Studies in Computational Intelligence 830

Maciej Huk
Marcin Maleszka
Edward Szczerbicki *Editors*

Intelligent Information and Database Systems: Recent Developments

Springer

Studies in Computational Intelligence

Volume 830

Series editor

Janusz Kacprzyk, Polish Academy of Sciences, Warsaw, Poland
e-mail: kacprzyk@ibspan.waw.pl

The series "Studies in Computational Intelligence" (SCI) publishes new developments and advances in the various areas of computational intelligence—quickly and with a high quality. The intent is to cover the theory, applications, and design methods of computational intelligence, as embedded in the fields of engineering, computer science, physics and life sciences, as well as the methodologies behind them. The series contains monographs, lecture notes and edited volumes in computational intelligence spanning the areas of neural networks, connectionist systems, genetic algorithms, evolutionary computation, artificial intelligence, cellular automata, self-organizing systems, soft computing, fuzzy systems, and hybrid intelligent systems. Of particular value to both the contributors and the readership are the short publication timeframe and the world-wide distribution, which enable both wide and rapid dissemination of research output.

The books of this series are submitted to indexing to Web of Science, EI-Compendex, DBLP, SCOPUS, Google Scholar and Springerlink.

More information about this series at http://www.springer.com/series/7092

Maciej Huk · Marcin Maleszka ·
Edward Szczerbicki
Editors

Intelligent Information and Database Systems: Recent Developments

Springer

Editors
Maciej Huk
Department of Information Systems
Wrocław University of Science and Technology
Wrocław, Poland

Marcin Maleszka
Department of Information Systems
Wrocław University of Science and Technology
Wrocław, Poland

Edward Szczerbicki
Department of Management
Gdańsk University of Technology
Gdańsk, Poland

ISSN 1860-949X ISSN 1860-9503 (electronic)
Studies in Computational Intelligence
ISBN 978-3-030-14131-8 ISBN 978-3-030-14132-5 (eBook)
https://doi.org/10.1007/978-3-030-14132-5

Library of Congress Control Number: 2019932697

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Program Committee

Ajith Abraham, Machine Intelligence Research Labs, USA
Muhammad Abulaish, South Asian University, India
Andy Adamatzky, University of the West of England, UK
Waseem Ahmad, Waiariki Institute of Technology, New Zealand
Bashar Al-Shboul, University of Jordan, Jordan
Lionel Amodeo, University of Technology of Troyes, France
Toni Anwar, Universiti Teknologi Malaysia, Malaysia
Tsuyoshi Arai, Okayama Prefectural University, Japan
Taha Arbaoui, University of Technology of Troyes, France
Ahmad Taher Azar, Benha University, Egypt
Thomas Bäck, Leiden University, The Netherlands
Amelia Badica, University of Craiova, Romania
Costin Badica, University of Craiova, Romania
Kambiz Badie, ICT Research Institute, Iran
Zbigniew Banaszak, Warsaw University of Technology, Poland
Dariusz Barbucha, Gdynia Maritime University, Poland
Ramazan Bayindir, Gazi University, Turkey
Juri Belikov, Tallinn University of Technology, Estonia
Maumita Bhattacharya, Charles Sturt University, Australia
Leon Bobrowski, Białystok University of Technology, Poland
Bülent Bolat, Yildiz Technical University, Turkey
Veera Boonjing, King Mongkut's Institute of Technology Ladkrabang, Thailand
Mariusz Boryczka, University of Silesia, Poland
Urszula Boryczka, University of Silesia, Poland
Zouhaier Brahmia, University of Sfax, Tunisia
Stephane Bressan, National University of Singapore
Peter Brida, University of Zilina, Slovakia
Andrej Brodnik, University of Ljubljana, Slovenia
Grażyna Brzykcy, Poznań University of Technology, Poland
Robert Burduk, Wrocław University of Science and Technology, Poland
Aleksander Byrski, AGH University of Science and Technology, Poland

David Camacho, Universidad Autonoma de Madrid, Spain
Tru Cao, Ho Chi Minh City University of Technology, Vietnam
Frantisek Capkovic, Institute of Informatics, Slovak Academy of Sciences, Slovakia
Oscar Castillo, Tijuana Institute of Technology, Mexico
Dariusz Ceglarek, Poznań High School of Banking, Poland
Zenon Chaczko, University of Technology, Sydney
Goutam Chakraborty, Iwate Prefectural University, Japan
Somchai Chatvichienchai, University of Nagasaki, Japan
Chun-Hao Chen, Tamkang University, Taiwan
Rung-Ching Chen, Chaoyang University of Technology, Taiwan
Shyi-Ming Chen, National Taiwan University of Science and Technology
Suphamit Chittayasothorn, King Mongkut's Institute of Technology Ladkrabang, Thailand
Sung-Bae Cho, Yonsei University, South Korea
Kazimierz Choroś, Wrocław University of Science and Technology, Poland
Kun-Ta Chuang, National Cheng Kung University, Taiwan
Dorian Cojocaru, University of Craiova, Romania
Jose Alfredo Ferreira Costa, Federal University of Rio Grande do Norte (UFRN), Brazil
Ireneusz Czarnowski, Gdynia Maritime University, Poland
Theophile Dagba, University of Abomey-Calavi, Benin
Quang A. Dang, Vietnam Academy of Science and Technology, Vietnam
Tien V. Do, Budapest University of Technology and Economics, Hungary
Grzegorz Dobrowolski, AGH University of Science and Technology, Poland
Habiba Drias, University of Science and Technology Houari Boumediene, Algeria
Maciej Drwal, Wrocław University of Science and Technology, Poland
Ewa Dudek-Dyduch, AGH University of Science and Technology, Poland
El-Sayed M. El-Alfy, King Fahd University of Petroleum and Minerals, Saudi Arabia
Keiichi Endo, Ehime University, Japan
Sebastian Ernst, AGH University of Science and Technology, Poland
Nadia Essoussi, University of Carthage, Tunisia
Rim Faiz, University of Carthage, Tunisia
Simon Fong, University of Macau, Macao
Dariusz Frejlichowski, West Pomeranian University of Technology, Szczecin, Poland
Takuya Fujihashi, Ehime University, Japan
Hamido Fujita, Iwate Prefectural University, Japan
Mohamed Gaber, Birmingham City University, UK
Ford Lumban Gaol, BINUS University, Indonesia
Dariusz Gąsior, Wrocław University of Science and Technology, Poland
Janusz Getta, University of Wollongong, Australia
Daniela Gifu, University Alexandru Ioan Cuza of Iasi, Romania
Dejan Gjorgjevikj, Ss. Cyril and Methodius University in Skopje, Macedonia

Barbara Gładysz, Wrocław University of Science and Technology, Poland
Daniela Godoy, ISISTAN Research Institute, Argentina
Ong Sing Goh, Universiti Teknikal Malaysia Melaka, Malaysia
Antonio Gonzalez-Pardo, Universidad Autonoma de Madrid, Spain
Yuichi Goto, Saitama University, Japan
Manuel Grana, University of Basque Country, Spain
Janis Grundspenkis, Riga Technical University, Latvia
Jeonghwan Gwak, Seoul National University, South Korea
Quang-Thuy Ha, VNU-University of Engineering and Technology, Vietnam
Dawit Haile, Addis Ababa University, Ethiopia
Pei-Yi Hao, National Kaohsiung University of Applied Sciences, Taiwan
Marcin Hernes, Wrocław University of Economics, Poland
Francisco Herrera, University of Granada, Spain
Kouichi Hirata, Kyushu Institute of Technology, Japan
Bogumiła Hnatkowska, Wrocław University of Science and Technology, Poland
Huu Hanh Hoang, Hue University, Vietnam
Quang Hoang, Hue University of Sciences, Vietnam
Van-Dung Hoang, Quang Binh University, Vietnam
Jaakko Hollmén, Aalto University, Finland
Tzung-Pei Hong, National Univesity of Kaohsiung, Taiwan
Mong-Fong Horng, National Kaohsiung University of Applied Sciences, Taiwan
Jen-Wei Huang, National Cheng Kung University, Taiwan
Yung-Fa Huang, Chaoyang University of Technology, Taiwan
Maciej Huk, Wrocław University of Science and Technology, Poland
Zbigniew Huzar, Wrocław University of Science and Technology, Poland
Dosam Hwang, Yeungnam University, South Korea
Roliana Ibrahim, Universiti Teknologi Malaysia, Malaysia
Dmitry Ignatov, National Research University Higher School of Economics, Russia
Lazaros Iliadis, Democritus University of Thrace, Greece
Hazra Imran, University of British Columbia, Canada
Agnieszka Indyka-Piasecka, Wrocław University of Science and Technology, Poland
Mirjana Ivanovic, University of Novi Sad, Serbia
Sanjay Jain, National University of Singapore, Singapore
Jarosław Jankowski, West Pomeranian University of Technology, Szczecin, Poland
Joanna Jędrzejowicz, University of Gdańsk, Poland
Piotr Jędrzejowicz, Gdynia Maritime University, Poland
Janusz Jeżewski, Institute of Medical Technology and Equipment ITAM, Poland
Geun Sik Jo, Inha University, South Korea
Kang-Hyun Jo, University of Ulsan, South Korea
Jason Jung, Chung-Ang University, South Korea
Przemysław Juszczuk, University of Economics in Katowice, Poland
Janusz Kacprzyk, Systems Research Institute, Polish Academy of Sciences, Poland
Tomasz Kajdanowicz, Wrocław University of Science and Technology, Poland
Nadjet Kamel, University Ferhat Abbes Setif1, Algeria

Hyun-Deok Kang, Ulsan National Institute of Science and Technology, South Korea
Mehmet Karaata, Kuwait University, Kuwait
Nikola Kasabov, Auckland University of Technology, New Zealand
Arkadiusz Kawa, Poznań University of Economics and Business, Poland
Rafał Kern, Wrocław University of Science and Technology, Poland
Zaheer Khan, University of the West of England, UK
Manish Khare, Dhirubhai Ambani Institute of Information and Communication Technology, India
Chonggun Kim, Yeungnam University, South Korea
Marek Kisiel-Dorohinicki, AGH University of Science and Technology, Poland
Attila Kiss, Eötvös Loránd University, Hungary
Jerzy Klamka, Silesian University of Technology, Poland
Goran Klepac, Raiffeisen Bank, Croatia
Shinya Kobayashi, Ehime University, Japan
Marek Kopel, Wrocław University of Science and Technology, Poland
Raymondus Kosala, BINUS University, Indonesia
Leszek Koszałka, Wrocław University of Science and Technology, Poland
Leszek Kotulski, AGH University of Science and Technology, Poland
Martin Kotyrba, University of Ostrava, Czech Republic
Jan Kozak, University of Economics in Katowice, Poland
Adrianna Kozierkiewicz, Wrocław University of Science and Technology, Poland
Bartosz Krawczyk, Virginia Commonwealth University, USA
Ondrej Krejcar, University of Hradec Kralove, Czech Republic
Dalia Kriksciuniene, Vilnius University, Lithuania
Dariusz Król, Wrocław University of Science and Technology, Poland
Marek Krótkiewicz, Wrocław University of Science and Technology, Poland
Marzena Kryszkiewicz, Warsaw University of Technology, Poland
Adam Krzyzak, Concordia University, Canada
Jan Kubíček, VSB—Technical University of Ostrava, Czech Republic
Tetsuji Kuboyama, Gakushuin University, Japan
Elżbieta Kukla, Wrocław University of Science and Technology, Poland
Julita Kulbacka, Wrocław Medical University, Poland
Marek Kulbacki, Polish-Japanese Academy of Information Technology, Poland
Kazuhiro Kuwabara, Ritsumeikan University, Japan
Halina Kwaśnicka, Wrocław University of Science and Technology, Poland
Annabel Latham, Manchester Metropolitan University, UK
Bac Le, University of Science, VNU-HCM, Vietnam
Hoai An Le Thi, University of Lorraine, France
Yue-Shi Lee, Ming Chuan University, Taiwan
Florin Leon, Gheorghe Asachi Technical University of Iasi, Romania
Horst Lichter, RWTH Aachen University, Germany
Kuo-Sui Lin, Aletheia University, Taiwan
Igor Litvinchev, Nuevo Leon State University, Mexico
Rey-Long Liu, Tzu Chi University, Taiwan

Doina Logofătu, Frankfurt University of Applied Sciences, Germany
Edwin Lughofer, Johannes Kepler University Linz, Austria
Lech Madeyski, Wrocław University of Science and Technology, Poland
Bernadetta Maleszka, Wrocław University of Science and Technology, Poland
Marcin Maleszka, Wrocław University of Science and Technology, Poland
Petra Maresova, University of Hradec Kralove, Czech Republic
Urszula Markowska-Kaczmar, Wrocław University of Science and Technology, Poland
Mustafa Mat Deris, Universiti Tun Hussein Onn Malaysia, Malaysia
Takashi Matsuhisa, Karelia Research Centre, Russian Academy of Science, Russia
Tamás Matuszka, Eötvös Loránd University, Hungary
João Mendes-Moreira, University of Porto, Portugal
Mercedes Merayo, Universidad Complutense de Madrid, Spain
Jacek Mercik, WSB University in Wrocław, Poland
Radosław Michalski, Wrocław University of Science and Technology, Poland
Peter Mikulecky, University of Hradec Kralove, Czech Republic
Marek Miłosz, Lublin University of Technology, Poland
Kazuo Misue, University of Tsukuba, Japan
Jolanta Mizera-Pietraszko, Opole University, Poland
Leo Mrsic, IN2data Ltd Data Science Company, Croatia
Agnieszka Mykowiecka, Institute of Computer Science, Polish Academy of Sciences, Poland
Paweł Myszkowski, Wrocław University of Science and Technology, Poland
Saeid Nahavandi, Deakin University, Australia
Kazumi Nakamatsu, University of Hyogo, Japan
Grzegorz J. Nalepa, AGH University of Science and Technology, Poland
Mahyuddin K. M. Nasution, Universitas Sumatera Utara, Indonesia
Richi Nayak, Queensland University of Technology, Australia
Fulufhelo Nelwamondo, Council for Scientific and Industrial Research, South Africa
Huu-Tuan Nguyen, Vietnam Maritime University, Vietnam
Le Minh Nguyen, Japan Advanced Institute of Science and Technology, Japan
Loan T. T. Nguyen, Nguyen Tat Thanh University, Vietnam
Ngoc-Thanh Nguyen, Wrocław University of Science and Technology, Poland
Quang-Vu Nguyen, Korea-Vietnam Friendship Information Technology College, Vietnam
Thai-Nghe Nguyen, Can Tho University, Vietnam
Van Du Nguyen, Wrocław University of Science and Technology, Poland
Yusuke Nojima, Osaka Prefecture University, Japan
Jerzy Nowacki, Polish-Japanese Academy of Information Technology, Poland
Agnieszka Nowak-Brzezińska, University of Silesia, Poland
Mariusz Nowostawski, Norwegian University of Science and Technology
Alberto Núñez, Universidad Complutense de Madrid, Spain
Manuel Núñez, Universidad Complutense de Madrid, Spain
Richard Jayadi Oentaryo, Singapore Management University, Singapore

Kouzou Ohara, Aoyama Gakuin University, Japan
Tarkko Oksala, Aalto University, Finland
Shingo Otsuka, Kanagawa Institute of Technology, Japan
Marcin Paprzycki, Systems Research Institute, Polish Academy of Sciences, Poland
Rafael Parpinelli, Santa Catarina State University (UDESC), Brazil
Danilo Pelusi, University of Teramo, Italy
Marek Penhaker, VSB—Technical University of Ostrava, Czech Republic
Hoang Pham, Rutgers University, USA
Maciej Piasecki, Wrocław University of Science and Technology, Poland
Bartłomiej Pierański, Poznań University of Economics and Business, Poland
Dariusz Pierzchała, Military University of Technology, Poland
Marcin Pietranik, Wrocław University of Science and Technology, Poland
Elias Pimenidis, University of the West of England, UK
Jaroslav Pokorný, Charles University in Prague, Czech Republic
Nikolaos Polatidis, University of Brighton, UK
Elvira Popescu, University of Craiova, Romania
Piotr Porwik, University of Silesia, Poland
Radu-Emil Precup, Politehnica University of Timisoara, Romania
Andrzej Przybyszewski, Polish-Japanese Academy of Information Technology, Poland
Małgorzata Przybyła-Kasperek, University of Silesia, Poland
Paulo Quaresma, Universidade de Evora, Portugal
Ngoc Quoc Ly, Ho Chi Minh City University of Science, Vietnam
David Ramsey, Wrocław University of Science and Technology, Poland
Mohammad Rashedur Rahman, North South University, Bangladesh
Ewa Ratajczak-Ropel, Gdynia Maritime University, Poland
Leszek Rutkowski, Częstochowa University of Technology, Poland
Tomasz M. Rutkowski, University of Tokyo, Japan
Henryk Rybiński, Warsaw University of Technology, Poland
Alexander Ryjov, Lomonosov Moscow State University, Russia
Keun Ho Ryu, Chungbuk National University, South Korea
Virgilijus Sakalauskas, Vilnius University, Lithuania
Daniel Sanchez, University of Granada, Spain
Rafał Scherer, Częstochowa University of Technology, Poland
Juergen Schmidhuber, Swiss AI Lab IDSIA, Switzerland
Jakub Segen, Polish-Japanese Academy of Information Technology, Poland
Ali Selamat, Universiti Teknologi Malaysia, Malaysia
S. M. N. Arosha Senanayake, Universiti Brunei Darussalam, Brunei Darussalam
Tegjyot Singh Sethi, University of Louisville, USA
Andrzej Siemiński, Wrocław University of Science and Technology, Poland
Dragan Simic, University of Novi Sad, Serbia
Paweł Sitek, Kielce University of Technology, Poland
Krzysztof Ślot, Łódź University of Technology, Poland
Adam Słowik, Koszalin University of Technology, Poland
Vladimir Sobeslav, University of Hradec Kralove, Czech Republic

Kulwadee Somboonviwat, Kasetsart University, Thailand
Kamran Soomro, University of the West of England, UK
Zenon A. Sosnowski, Białystok University of Technology, Poland
Bela Stantic, Griffith University, Australia
Stanimir Stoyanov, University of Plovdiv “Paisii Hilendarski”, Bulgaria
Ja-Hwung Su, Cheng Shiu University, Taiwan
Libuse Svobodova, University of Hradec Kralove, Czech Republic
Jerzy Świątek, Wrocław University of Science and Technology, Poland
Andrzej Świerniak, Silesian University of Technology, Poland
Tadeusz Szuba, AGH University of Science and Technology, Poland
Julian Szymański, Gdańsk University of Technology, Poland
Ryszard Tadeusiewicz, AGH University of Science and Technology, Poland
Yasufumi Takama, Tokyo Metropolitan University, Japan
Maryam Tayefeh Mahmoudi, ICT Research Institute, Iran
Zbigniew Telec, Wrocław University of Science and Technology, Poland
Aleksei Tepljakov, Tallinn University of Technology, Estonia
Dilhan Thilakarathne, Vrije Universiteit Amsterdam, Netherlands
Satoshi Tojo, Japan Advanced Institute of Science and Technology, Japan
Krzysztof Tokarz, Silesian University of Technology, Poland
Bogdan Trawiński, Wrocław University of Science and Technology, Poland
Krzysztof Trojanowski, Cardinal Stefan Wyszyński University in Warsaw, Poland
Ualsher Tukeyev, Al-Farabi Kazakh National University, Kazakhstan
Aysegul Ucar, Firat University, Turkey
Olgierd Unold, Wrocław University of Science and Technology, Poland
Joost Vennekens, Katholieke Universiteit Leuven, Belgium
Jørgen Villadsen, Technical University of Denmark, Denmark
Bay Vo, Ho Chi Minh City University of Technology, Vietnam
Eva Volna, University of Ostrava, Czech Republic
Wahyono, Universitas Gadjah Mada, Indonesia
Lipo Wang, Nanyang Technological University, Singapore
Junzo Watada, Waseda University, Japan
Konrad Wojciechowski, Silesian University of Technology, Poland
Krystian Wojtkiewicz, Wrocław University of Science and Technology, Poland
Krzysztof Wróbel, University of Silesia, Poland
Marian Wysocki, Rzeszow University of Technology, Poland
Farouk Yalaoui, University of Technology of Troyes, France
Xin-She Yang, Middlesex University, UK
Tulay Yildirim, Yildiz Technical University, Turkey
Piotr Zabawa, Cracow University of Technology, Poland
Sławomir Zadrożny, Systems Research Institute, Polish Academy of Sciences, Poland
Drago Žagar, University of Osijek, Croatia
Danuta Zakrzewska, Łódź University of Technology, Poland
Constantin-Bala Zamfirescu, Lucian Blaga University of Sibiu, Romania
Katerina Zdravkova, Ss. Cyril and Methodius University in Skopje, Macedonia

Vesna Zeljkovic, Lincoln University, UK
Aleksander Zgrzywa, Wrocław University of Science and Technology, Poland
Qiang Zhang, Dalian University, China
Maciej Zięba, Wrocław University of Science and Technology, Poland
Adam Ziębiński, Silesian University of Technology, Poland
Beata Zielosko, University of Silesia, Poland
Marta Zorrilla, University of Cantabria, Spain

Preface

The ongoing integration of various intelligent information, database systems and technologies—including artificial intelligence, social media, multimedia, cloud computing, and big data processing—leads to important synergy of those fields. It results in new practical solutions and applications as well as broadens the scope of theoretical analyses and approaches. But it also creates previously not considered concerns and problems, related not only to the scale but also to new functionalities of created methods and systems. This includes such issues as reliability and long-term integrity of critical applications, energy efficiency of big data storage, processing, organization, and quality assurance during the development of very complex systems. Continuous research in the fields of intelligent information and database systems provides an opportunity to identify and address those problems by suggesting and creating further solutions and technologies.

The main goal of this book is to present and intensify practical application of current knowledge and state of the art emerging in the field of intelligent information and database systems. Thus, it includes a set of carefully selected research papers presenting new methods, technologies and applications. The represented scope of topics encompasses a wide spectrum of research subjects discussed both from the theoretical and practical points of view. The book covers such fundamental aspects of intelligent information and database systems as artificial and computational intelligence, nature-inspired algorithms and paradigms, collective knowledge and ontologies, computer vision techniques, and multi-dimensional large-scale data processing. At the same time, it presents a number of interdisciplinary topics including social networks, system security, cloud computing, Internet of things, business intelligence, and biomedical data analysis.

This volume in the well-established "Studies in Computational Intelligence" series can be a valuable source of knowledge about current problems and solutions in the field. It contains a selection of 34 chapters based on original research presented as posters during *the Asian Conference on Intelligent Information and Database Systems* (*ACIIDS 2019*) held on 8–11 April 2019 in Yogyakarta, Indonesia. This eleventh edition of the ACIIDS conference was organized by the

Wrocław University of Science and Technology, Poland, together with BINUS University, Indonesia, and their partners.

The book will be a useful resource and reference for researchers and practitioners interested in increasing synergy between artificial intelligence and database technologies as well as for graduate and Ph.D. students in computer science and related fields.

Finally, on behalf of the Steering Committee, Program Committee and Organizing Committee of the ACIIDS 2019 we would like to thank all the authors for sharing their research during the conference. We also extend our thanks to all reviewers for their time, engagement, and expertise. And special gratitude we express to Prof. Janusz Kacprzyk, the editor of this series, and Dr. Thomas Ditzinger from Springer for their interest and support for our project. The help of those people was invaluable in the preparation of this book.

Yogyakarta, Indonesia
April 2019

Maciej Huk
Marcin Maleszka
Edward Szczerbicki

Contents

Part I
Sensor Clouds and Internet of Things

Proposing Smart Disaster Management in Urban Area

Dyah Wahyu Sukmaningsih, Wayan Suparta, Agung Trisetyarso, Bahtiar Saleh Abbas and Chul Ho Kang

Abstract This study identifies technology and data related to disaster management with the concept of the Smart City. A natural disaster has a huge impact on city sustainability, it could destroy physical infrastructure, harm people, or make activities stop for a range of time. Environment evolvement in recent years, such as climate change and increasing population impacted to the natural disaster casualty. It then examines the technology to construct natural disaster management, consisting of preparation, mitigation, response and recovery for disaster recovery. Technology, such as IoT, GIS, big data, AUV, and Social media can be integrated to provide a solution for the city. Furthermore, this study proposed a model for smart disaster management that integrates ICT (Information and Communications Technology) and Smart City domain.

Keywords Smart city · Disaster management · Natural disasters

1 Introduction

Currently, 55% of the world's population lives in urban areas, and this proportion is predicted to have 68% growth by 2050. Urbanization, which means a shift in

D. W. Sukmaningsih (✉) · W. Suparta · A. Trisetyarso · B. S. Abbas · C. H. Kang
Computer Science Department, BINUS Graduate Program-Doctor of Computer Science, BINUS University, Jakarta 11480, Indonesia
e-mail: dyah.wahyu@binus.ac.id

W. Suparta
e-mail: drwaynesparta@gmail.com

A. Trisetyarso
e-mail: atrisetyarso@binus.edu

B. S. Abbas
e-mail: bahtiars@binus.edu

C. H. Kang
e-mail: chkang5136@kw.ac.kr

M. Huk et al. (eds.), *Intelligent Information and Database Systems: Recent Developments*, Studies in Computational Intelligence 830,
https://doi.org/10.1007/978-3-030-14132-5_1

the population from rural to urban areas related with world population growth, an additional 2.5 billion people will live in urban areas by 2050, nearly 90% of them in Asia and Africa. Many countries will have to face the challenges of their growing urban populations, that includes some consideration about transport, housing, energy systems, and infrastructure, as well as education, employment, health care, decent work and a safe environment [1]. When the city becomes more important in the future, a smart city is a concept to utilize ICT to help the citizens in a very day life and make the city more sustainable.

A disaster is a phenomenon that causes tremendous losses of life, property and the environment in a very short period of time. Nowadays when urbanization is at its peak, most of the population density would be concentrated on the city. Therefore, disasters that occur in urban areas with such a large population can cause the death of many people and loss of property [2]. To best of our knowledge, disaster management is rarely discussed as part of smart city initiative. Despite the fact that for any event of a disaster will affect city sustainability. For practical reason, this paper will examine only natural disaster, which is a disaster that caused by the mother nature. The aim of the present research is to identify what kind of technology that can be integrated into smart disaster management and which information and data will be provided.

2 Literature Review

2.1 Smart City

To get insight about the concept of the smart city, we need to understand why cities have a significant and important role in the future, especially in the social and economic aspects of the world, and that they also have a strong impact on the environment. A smart city definition has been reviewed by [3] and [4], keep pace with changes in the field of information technology as a new approach to mitigate and give solutions for the urban problem and make development more sustainable.

There also two domains that applied in this term. First, the domain where ICT becomes a machine used in assisting cities in overcoming problems, for example, energy grids, water management, natural resources, mobility, waste management, and logistic. Secondly, this smart city has also been applied to "soft domain" such as culture, education, social inclusion, government, and political innovation, where the use of ICT is usually not too crucial, but as additional support. Neirotti classified each domain and sub-domains in the smart city as described in Fig. 1 [5].

As state by Cocchia [4], smart city definition by Giffinger and Gudrun [6] is the most cited, "A Smart City is a city well performing built on the 'smart' combination of endowments and activities of self-decisive, independent and aware citizens". Albino et al. [3] has been trying to provide some of the key factors and concepts of Smart City, therefore Smart City's focus is not just on smart technology, but also on solving core urban and regional issues. There are three fundamental issues that need to be

Fig. 1 Domain and sub-domain in smart city

addressed in cities whose smart technology can help solve these problems: culture, metabolism and governance.

Ojo et al. [7] summarize the concept of "smart city" as an interaction between technological innovation, organizational innovation, and policy innovation. Sectors that are typically affected by smart city initiatives include transport, environment, governance, energy, lifestyle and people, technology and infrastructure.

Significant progress has been made in information technology, such as sensors, actuators, and smartphones, which create great business potential for the Internet of Things, as all devices can be interconnected and communicate on the Internet. Then data can be analyzed using big data technology.

2.2 Disaster Management

This paper will take a definition of disaster from two UN agencies, CRED (Centre for Research on the Epidemiology of Disaster) and WHO (World Health Organization). CRED defines a disaster as "a situation or event which overwhelms local capacity, necessitating a request at national or international level for external assistance; an unforeseen and often sudden event that causes great damage, destruction and human suffering". The WHO definition of natural disaster state that "A natural disaster is an act of nature of such magnitude as to create a catastrophic situation in which the day-to-day patterns of life are suddenly disrupted and people are plunged into helplessness and suffering, and, as a result, need food, clothing, shelter, medical and nursing care and other necessities of life, and protection against unfavourable environmental factors and conditions."

United Nation classifies natural disaster to Floods, landslides, earthquakes, extreme temperature, storms, wildfire, volcanic activity, drought, and mass movement-dry. Disasters occur when natural hazards that lead to the death or accident of people who are at risk or the loss of valuables. Vulnerability emerges and increases for various reasons, including population growth, urban development at vulnerable sites, land use change, environmental degradation, weak governance, poverty and inequality, and climate change [8]. UNISDR report, that between 1998 and 2017, climate-related and geophysical catastrophes killed 1.3 million people and caused another 4.4 billion injuries, homeless, displacement or emergency relief.

In managing disaster, it is necessary to understand the Disaster Management phase, therefore solutions have to related to the following phase: preparation, response, recovery, and mitigation. (1) Hazard mitigation is a continuous effort to reduce the physical and social impact of future disasters. There are two main types of mitigation: Structural mitigation and non-structural mitigation (2) preparedness, action taken before a disaster to deal with anticipated disaster response and recovery. (3) Response, Direct protection of life and property, damage control and minimizing the impact of disasters. (4) Recovery, in this case both short-term activities, to restore the vital physical and social systems, concurrently for long-term activities aimed at restoring system to pre-disaster conditions [9, 10].

ICT (Information Communication Technology) has become an indispensable part of implementing disaster management. The application of ICT at the beginning had the effect of weather forecast and communication for disaster announcements. The advance of the new technology can provide an improvement in the field of disaster management.

2.3 Technology Solution

Smart disaster system task is to perform any action pre-disaster until post-disaster that give warrants to save people and protect property. To properly manage disaster events, information must be collected, such as: by sharing resources (data or information) and coordinating, making decisions and activities. In addition, these resources and data must be collaborated to accomplish complex tasks such as evacuation of geographic areas and operations with actuators. The insufficiency of integrated platforms and infrastructures leads to failure of emergency management [11].

IoT (Internet of Things)

IoT is a broadband network that uses a standard communication protocol whose point of convergence is the Internet. The main concept of IoT is the presence of universal objects that can be measured, derived and understood, and that can transform the environment. Because of this, IoT is made possible by the development of various objects and communication technologies. Example of the IoT technology found in smart devices, including cell phones and other items such as an appliance, landmarks, monuments, and works of art that can be general objectives [12].

Using IoT, system can integrate between internet and any device/object. Sensors can be developed at different location for any purposes (e.g. Weather and water system, environmental pollution, surveillance system). IoT in smart disaster system that installs in Weather and water systems can apply some sensors to provide appropriate information like temperature, wind speed, rain, and pressure.

Cloud-Based Big Data

Big data can also activate disaster resilience for smart cities. The use of huge data and data analytics in smart city development is increasing resulting from intelligence processing, exploitation, dissemination, and analysis. Today we see unusual interdependence on a local and global scale, and today's risks are systemic, viral, devastating and unpredictable consequences. Therefore, traditional emergency management methods or technology cannot encounter this new challenge. Big data, with the advantages of data collection, data merging, data storage, data analysis, data representation, and data visualization can be used to mine patterns, analyze user behavior, visualize, and track data among communities to overcome the underlying vulnerability. With big data, disaster managers can identify potential risks, make actions to prevent risks and ultimately enable the resilience of the city [13].

The development of crisis analysis-based on the big data platform offers many applications that can improve the response to disasters. Disasters often have a sequential impact (e.g. tsunamis followed by earthquakes). Analysis platforms have to be developed in real time to prioritize urgent problems, avoid further dangers, for example, coordinate aid organizations and organize first responders. Some features that can be done using big data include: Emergency services need to process large amounts of data that are entered through multiple channels, such as first responders, satellite networks, sensors, and social media. Therefore, the major challenge is to develop data integration protocols that ensure proper data collection, modeling, and messaging systems [10].

GIS (Geographic Information System) and Remote Sensing

Natural disaster management requires many types of information that have a spatial component. Geospatial data includes data with geographical components such as maps, aerial photographs, GPS data, rainfall data, satellite imagery, Roads, land use, hydrography, elevation, boundaries, image base, etc. Information from remote sensing and GIS has proven useful in disaster management.

Remote sensing is an activity that uses satellite or sensor from aircraft to detect and classify objects on earth. Remote sensing is one of the most advanced ways to detect forest fires. Spaceborne (satellite) instruments with built-in sensors provide information about forest fires on the earth's surface. An example of remote sensing is the use of lasers to tracking smoke plume boundaries, its top height, dispersion, and a change of its intensity with time [14], another example of remote sensing usability is to detect landslides [15]. Westen [16] describes GIS and remote sensing role in disaster management: marking left by previous disasters, recognize the initial stage of the event, monitoring events, planning evacuation route, search and rescue operation.

Data integration for GIS and types of information are required [16, 17]:

- Data at the disaster phenomenon (e.g. floods, Landslides, earthquakes), their frequency, vicinity, price, etc
- Data at the environment where the disaster possibly take regions: geology, topography, hydrology, geomorphology, land use, soils, plant life and masses of others.
- Data on the factors that are likely destroyed if the event takes place: settlements, infrastructure, the populace, socio-monetary statistics and so on.

- Data at the emergency comfort property, which embody hospitals, police stations, hearth brigades, warehouses and loads of others.

Social Media
Nowadays, social media emerged as an important party and playing a significant role in disaster management and response. Social media has a potential source of information to improve the management of emergency situations such as disasters triggered by natural hazards. Recent research has concentrated on using social media not as a stand-alone information source but also join with other information. Social media has advantages include crowd-sourcing, speed, and easy to access from mobile devices. It also has disadvantages include bias in user base; confusing information between related and informative and false content, and out-of-date information. Twitter is one of the most used social media in term of disaster information, GPS in the smartphone provides geo-tagged for twitter users [18, 19].

UAV (Unmanned Aerial Vehicle)/Drones
UAV (Unmanned aerial vehicle) or drone playing a vital role as first responders during the disaster. UAV can achieve a better understanding of which infrastructure are affected by the event, how much this infrastructure is damaged, how the traffic infrastructure is affected and how many potential people are affected by the event [20].

Some application of UAV in the disaster event such as Monitoring and early warning, disaster information sharing, evacuation support, communication system, SAR, Damage assessment.

3 Research Method

This paper proposes the model for mapping disaster management and smart city domain. Following this model, apply technology related to disaster management solution. This paper examines from a study that has ever existed, then put technological solutions related to the domains in smart cities that are connected to disaster management.

For the purpose of understanding how social media used during the rainy season, this paper examines twitter usage monitor tweet related to "banjir" (flood) in Indonesia. Flooding is an event that often occurs during the rainy season in Indonesia. The event not only happened because of high rainfall but also because of human factors and infrastructure. The peak of a rainy season usually happened in January until February, but the data gathered from early December when rainy season has just begun.

The data were collected by querying using NodeXL (basic) and Twitter API from 28 November 2018 to 2 December 2018, this collection contains on average about 1262 tweets and mention, and data was sampled by tweet every day after 06.00 p.m. We further filtered tweets by keyword "banjir".

But, to request tweet with a geolocation, that indicates user location, there is difficulty that NodeXL follows rule from twitter search and location search was deactivate. Consequently, for analyzing user location to match with tweets in urban area, we use manual searching for city in Indonesia, like "Jakarta", "surabaya" and other location that indicates place in city area like "fatmawati" which is sub-district in Jakarta.

Afterward, we further cleansing tweet and excluded those outside Indonesia and "mentions" and "replies to" type of tweet. Total tweet collected are 290, tweet then classified based on the thematic coding, tweet coded manually and group into seven thematic group adapted categories from [18]. Group of thematic categories consist of (1) 'volunteer actions', (2) 'media reports', (3) 'traffic conditions', (4), 'first-hand observations', (5) 'official actions', (6) 'infrastructure damage' and (7) 'other'. Table 4 present result from tweet thematic classification.

This study also classifies tweet by its content, using classification adapted from [21], these categories are related-informative, related-not informative and not related. Related category concern about tweets that talks "banjir" as flood event no other meaning of the word. For example "banjir itu berkah" (the flood is a blessing) was categorize as not related. Informative category use for tweets that gives useful information to flood event and not for other purposes. For example "yg mau melintasi jalur purwodadi pati harap berhati, Banjir yg arusnya kuat dari atas" (who want to cross the purwodadi pati, please be careful) categorize into related-informative. Tweet categorization based on relatedness presented in Table 3.

4 Results and Discussion

4.1 Smart Disaster Management

A sustainable city is inseparable with managing disaster event. Combining possible technology will make it more effective and useful for both the city and its citizen. Figure 2 describes a technology which already uses in disaster management. The model integrates technology with smart city domain with the intention to give more quick and reliable solution for the future event.

An action was taken in every phase of disaster management that state by National Research Council (NRC) [9] mapping into smart city domain described by Neirotti et al. [5] to give clear visibility for authorities and citizen. Action and data provide from disaster management described in Table 1, every phase is fulfilled by smart city domain and contribution that smart city should plan in disaster management.

Furthermore, technologies solution for smart disaster management describe in Table 2. In addition, these technologies will provide information which available for authorities to use and optimize to gain advantage from smart disaster system presented. This solution was obtained from previous literature and research, for instance [16] already presented the utilization of GIS and remote sensing for every phase of

Table 1 Mapping disaster phase and smart city domain

Smart city domain	Disaster phase			
	Preparedness	Mitigation	Response	Recovery
City logistics	Mapping logistic route	Logistic evacuation support	Send logistic along with the rescue team	
People mobility	Mapping evacuation route	• Evacuate geographical area • Tracking people location	• Alarm people through mobile phone • Prediction population displacement • Identify area affected	
Pollution control	Install IoT in water and weather system	• Analyze data from remote sensing • Analyze anomaly of precursor	• Spatial monitoring	
Public safety	• Potential risk identification • Identify important place for people	• Crowd-sourcing for disaster information • Early warning	• Sending search and rescue team	• Predicting future event and impact
Health care		Mapping rescue team route	• Dispatch emergency and rescue team • Disaster response system	
Building	Mapping danger zone			• Damage assessment for reconstruction • Update danger zone

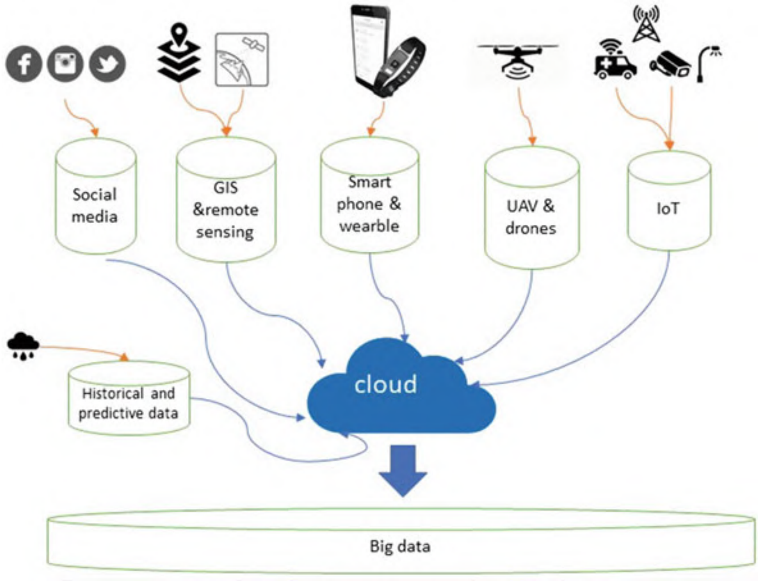


Fig. 2 Smart disaster management

disaster management. Recent technology introduced UAV and drones for image capture and aerial data gathering, using UAVs can make observations of disaster-affected areas faster and real-time [20].

Finally big data consolidate all the data and analyze them then provide valuable information for agency, government and emergency services personnel. With the aid of these information, it is expected that decision making can be done more accurately. In some cases, there is a combination and integration between each technology, data from different platform has enormous amount of specific data (location, sensor, spatial), big data can provide real-time information for emergency services.

The flood forecasting system design for the Krishna River valley with GIS and meteorological data. The map is created with ArcGIS, which includes the depth of inundation and the degree of flooding with presentation maps. This information will be available to the relevant authorities in the near future in order to take the necessary action [22].

Recent experience shows that the public can use social media and Web 2.0 interactive Web sites to communicate effectively and inform online colleagues about event location information by creating map-based mashups. Cartographer Experts and emergency response professionals can benefit and collect various types of information [23].

Study on Floods in Germany and how related information from Twitter and real-time event show significantly information from Twitter as disaster response and monitoring management. From that study found that Tweets related to the high of water level related to the proximity of the tweet. The study correlates the proximity of

Table 2 Technology applied for smart disaster management

Technology	Applied	Data
IoT	• IoT sensor on water gauge and weather system • IoT sensor for temperature, CO and CO_2 level in the forest for monitoring forest fire • IoT on the ocean for detecting a seismic wave (e.g. Tsunami)	Water sensor, weather sensor, emergency service wireless network, surveillance, seismic, air pollution
GIS and smart sensing	• Provide spatial data of the affected area (mapping disaster area) • Zoning the area based on risk magnitude • Simulating damage scenario • Planning evacuation and operation • Damage assessments for rehabilitation and reconstruction	Spatial data (maps, satellite imagery, aerial photographs, GPS data, Roads, land use, boundaries, rainfall data, hydrography, elevation)
Social media	• Communication and interaction between agency and community • Information dissemination • Early warning	Geo-location, hashtag, frequency of tweets
Smartphone and wearable	• Early warning • Tracking, navigating searching and rescuing	GPS
UAV/drones	• First responder, capture area affected by the disaster • Access to hard-to-reach-area	Image, GPS/location
Big data	• Analyze	

tweet using geo-tagged and water gauge sensor [18]. Another work [19] examining twitter usage in Padang, suggests that government adopts Twitter as an element of disaster management for planning and an early warning.

Table 3 shows the distribution of relatedness found in flood event, 43% of the tweet is not related to "banjir", and only 27% of the tweets related and informative, Table 4 shows the distribution of content-type found in flood event. "other" content type has the highest portion of 64% of tweet, with some portion of the tweets is running by tweet bot. The second highest portion 21% are tweets from online media, reporting about flood event. From this data, we can observe considerable tweets that are not informative and most of the informative tweet is from online media. But this can also occur because the rainy season is not yet at its peak, so people have not felt

Table 3 Tweets categorized by relatedness

Period	28 Nov 2018		29 Nov 2018		30 Nov 2018		01 Dec 2018		2 Dec 2018		Total	
All Tweets	27	100%	21	100%	58	100%	90	100%	94	100%	290	100%
Related-informative	7	26%	4	19%	12	21%	30	33%	25	27%	78	27%
Related not informative	9	33%	10	48%	15	26%	13	14%	40	43%	87	30%
Not related	11	41%	7	33%	31	53%	47	52%	29	31%	125	43%

Table 4 Tweets categorized by content type

Period	28-Nov-18		29-Nov-18		30-Nov-18		01 Dec 2018		02 Dec 2018		Total	
All Tweets	27	100%	21	100%	58	100%	90	100%	94	100%	290	100%
Volunteer action	0	0%	0	0%	0	0%	2	2%	0	0%	2	1%
Media report	10	37%	7	33%	17	29%	31	34%	20	21%	85	29%
First observation	0	0%	4	19%	3	5%	1	1%	0	0%	8	3%
Official action	2	4%	0	0%	1	2%	2	2%	3	3%	7	3%
Infrastructure damage	0	0%	0	0%	0	0%	1	1%	0	0%	1	0%
Traffic condition	0	0%	0	0%	1	2%	0	0%	0	0%	1	0%
Other	15	56%	10	48%	36	62%	53	59%	71	76%	185	64%

a crisis. Caution and advice tweet about flood situation frequently came from online media.

On the search for location related tweet, we found that some of tweets that indicate for location, for example

"Wah musim hujan jakarta banjir lagi nih" (Wow, the Jakarta rainy season floods again)
"Warga Bekasi Berharap Solusi Untuk Atasi Banjir" (Bekasi Residents Hoping for Solutions to Overcome Floods)

Some of the locations that mention in tweets related to location are: Jakarta/DKI, Bekasi, Jembarna (Bali), Pesawaran (Lampung), Bandung, Fatmawati (Jakarta sun-district), Tangerang, Gumilir (cilacap), Sunter (Jakarta sub-district).

Most of the tweet from residents that mention about location, is mention city area. Tweet from media and government covers a wider area like province and village. This is an interesting insight if there is a disaster that has a greater impact, then it is probable that the information from Twitter will come from urban areas.

Do not conclude that information from social media is useless, we can position twitter media to provide insight if we can understand the details of diffusion from the twitter data. Government or any disaster agency who need to develop disaster system must identify data from Twitter that is relevant to the disaster event. and seeing that most tweets come from urban areas, this can be a consideration for using this information as a warning system in mitigation in disaster situations and apply this to the smart city system.

5 Conclusions

Smart disaster management models in this document provide insight into technologies that can help with disaster management, and how they can be integrated into a smart city model. As a result, it can provide up-to-date information and references to take action that can prevent and minimize disasters. Furthermore, in the event of the disaster, the system is able to display real-time situation and execute a rescue plan.

From the previous study, there is some advanced technology that useful for providing solutions in disaster management. These technologies provide data and information that can be used to support actions in each phase in disaster management.

Analysis from twitter data shows the participation of the community to provide information. Active social media users are widely located in urban areas, information from these users provide information if analyze more thoroughly. The government must take steps and find ways to optimize the benefits of social media. Noticing the effectiveness of users on social media, the government must also be proactive, not only gathering information but also using social media as a tool to communicate accurate information to citizens.

To capitalize on the potential of implementing intelligent disaster management, the future work has to explore any potential collaboration between other technology. For example, information from social media integrates with GIS and remote sensing, or collaboration between UAV and social media using semantic web and machine learning to analyze data.

References

1. United Nation: 68% of the world population projected to live in urban areas by 2050. https://www.un.org/development/desa/en/news/population/2018-revision-of-world-urbanization-prospects.html
2. Arepalli, A., Rao, S.S., Rao, P.J.: A spatial disaster management framework for smart cities—a case. In: Proceedings of International Conference on Remote Sensing for Disaster Management: Issues and Challenges in Disaster Management, p. 465 (2018)
3. Albino, V., Berardi, U., Dangelico, R.M.: Smart cities: definitions, dimensions, performance, and initiatives. J. Urban Technol. **22**, 1–19 (2015)
4. Cocchia, A.: Smart and digital city: a systematic literature review. In: Dameri, R.P., Rosenthal-Sabroux, C. (eds.) Smart city: how to create public and economic value with high technology in urban space, pp. 13–43. Springer International Publishing, Cham (2014)
5. Neirotti, P., De Marco, A., Cagliano, A.C., Mangano, G., Scorrano, F.: Current trends in smart city initiatives: some stylised facts (2014)
6. Giffinger, R., Gudrun, H.: Smart cities ranking: an effective instrument for the positioning of the cities? ACE Archit. City Environ. **4**, 7–26 (2010)
7. Ojo, A., Curry, E., Zeleti, F.A.: A tale of open data innovations in five smart cities. In: Proceedings of the Annual Hawaii International Conference on System Sciences, 2015 March, pp. 2326–2335 (2015)
8. UNISDR: Economic losses, poverty & disasters 1998–2017 (2017)
9. National Research Council: Improving disaster management: the role of IT in mitigation, preparedness, response, and recovery. National Academies Press (2007)
10. Akter, S., Wamba, S.F.: Big data and disaster management: a systematic review and agenda for future research. Ann. Oper. Res. 1–21 (2017)
11. Sakhardande, P., Hanagal, S., Kulkarni, S.: Design of disaster management system using IoT based interconnected network with smart city monitoring. In: 2016 International Conference on Internet of Things and Applications IOTA 2016, pp. 185–190 (2016)
12. Arasteh, H., Hosseinnezhad, V., Loia, V., Tommasetti, A., Troisi, O., Shafie-khah, M., Siano, P.: IoT-based smart cities: a survey IoT-based smart cities: a survey. In: 2016 IEEE 16th International Conference on Environment and Electrical Engineering (EEEIC), pp. 2–7 (2016)
13. Yang, C., Su, G., Chen, J.: Using big data to enhance crisis response and disaster resilience for a smart city. In: 2017 IEEE 2nd International Conference on Big Data Analytics ICBDA 2017, pp. 504–507 (2017)
14. Meenakshi, Y., Kumar, Y.B.: Multi-angle LIDAR for remote sensing smoke emissions from wildfires. In: Proceedings of International Conference on Remote Sensing for Disaster Management, pp. 83–89 (2019)
15. Shaik, I., Rao, S.V.C.K., Penta, B.: Detection of landslide using high-resolution satellite data and analysis using entropy. In: Proceedings of International Conference on Remote Sensing for Disaster Management, pp. 243–250 (2019)
16. Westen, C.V.A.N.: Remote sensing for natural disaster management. **XXXIII**, 1609–1617 (2000)
17. Anjaneyulu, G., Suseela, A.: Role of remote sensing system for disaster area response. In: Proceedings of International Conference on Remote Sensing for Disaster Management, pp. 151–157 (2019)
18. de Albuquerque, J.P., Herfort, B., Brenning, A., Zipf, A.: A geographic approach for combining social media and authoritative data towards identifying useful information for disaster management. Int. J. Geogr. Inf. Sci. **29**, 667–689 (2015)
19. Carley, K.M., Malik, M., Landwehr, P.M., Pfeffer, J., Kowalchuck, M.: Crowd sourcing disaster management : the complex nature of Twitter usage in Padang Indonesia. Saf. Sci. (2016)
20. Erdelj, M., Natalizio, E., Chowdhury, K.R., Akyildiz, I.F.: Help from the sky: leveraging UAVs for disaster management, 24–32 (2017)

21. Olteanu, A., Vieweg, S., Castillo, C.: What to expect when the unexpected happens: social media communication across crises. In: Proceedings of 18th ACM Conference on Computer-Supported Cooperative Work and Social Computing—CSCW'15, pp. 994–1009 (2015)
22. Duvvuri, S.: GIS-based management system for flood forecast applications. In: Proceedings of International Conference on Remote Sensing for Disaster Management, pp. 1–12 (2019)
23. Tao, W.: Interdisciplinary urban GIS for smart cities: advancements and opportunities. Geo-Spat. Inf. Sci. **16**, 25–34 (2013)

Ubiquitous Rehabilitation Combining Inertial Measurement System with Smartphone and Supported by Visual and Voice Feedback

Bartłomiej Pędrys, Henryk Josiński and Konrad Wojciechowski

Abstract The positive influence of physical activity on people's health and well-being is no longer a doubt. Instead of "do we need to exercise?" people ask, "how to exercise?" what is noticeable in growing interest in active lifestyle websites. Following the trend, additionally considering the popularity of smartphone applications, affordability and still decreasing size of inertial measurement unit based motion capture solutions, a new system prototype has been proposed. Smartphone connected with inertial measurement sensors will allow users to perform an independent, ubiquitous rehabilitation.

Keywords Ubiquitous rehabilitation · Inertial measurement unit · Smartphone · Attitude and heading reference system · Unity

1 Introduction and Motivation

At the beginning, it is crucial to discuss three important aspects. First, with the constant development of technology, people got used to still increasing ease of performing the same actions e.g. simplicity of taking picture with camera embedded into the phone which superseded photographic processing. Whenever possible, multiple tasks are done simultaneously. Smartphone is a type of solution for such need, and that is the reason why it attracts so many users.

The second aspect relates to increasing trend of sedentary lifestyle. Office based jobs connected with no physical activity have a negative influence on people's health. Even the simplest exercises can significantly improve blood circulation what results in better oxidation of the brain. This type of exercises may be useful in the process of rehabilitation e.g. for people recovering from surgeries or car accidents.

B. Pędrys (✉) · H. Josiński · K. Wojciechowski
Polish-Japanese Academy of Information Technology, Koszykowa 86,
02-008 Warsaw, Poland
e-mail: bart.pedrys@gmail.com

M. Huk et al. (eds.), *Intelligent Information and Database Systems: Recent Developments*, Studies in Computational Intelligence 830,
https://doi.org/10.1007/978-3-030-14132-5_2

The last aspect is the increasing affordability of inertial measurement sensors combined with their decreasing dimensions. IMUs (inertial measurement sensors) will soon be classified as wearable hardware what will be a beginning of their everyday use.

A system that combines the aforementioned aspects was designed to create a new product that responds to the problems of today's community. The paper describes the stage of planning and implementation of the system's prototype, and its tests.

2 Background

Popularity of rehabilitation aided by computer visualisation is growing. Beside custom systems developed for such applications, there are many other devices e.g. Kinect, IMU-based suits, or smartphones with suitable sensors integrated [1, 2, 5]. Commonness of devices which involve mentioned sensors (e.g. smartphones) makes these solutions much more affordable.

Mobile devices which are present in almost everyone's life can aggregate data about their activity. Accurate analysis of gathered big data can result in targeted diagnoses e.g. faulty posture diagnose [4]. Current computing power combined with popularity of smartphones caused more frequent use of them during rehabilitations [7].

Authors of newest rehabilitation applications are using as advanced solutions as possible to make this form more user-friendly [8]. Currently the attention is focused on creating easy to interpret forms of graphical representation of collected data [9]. Especial example is a 3D avatar which is an exact projection of the patient [10, 11].

One of the most popular forms of entertainment are games for mobile devices. This fact can be used to make rehabilitation processes more attractive [6, 12]. In addition, smartphones are able to send notifications which can help patient exercise regularly.

Except easy to interpret visualisations, one of the most important aspects are mobility of the system and its low cost [5].

Innovation of presented solution is a combination of all mentioned advantages. Special attention has been paid to: low cost of the device, high popularity of target equipment, easy interpreted visualisations and the intuitive control. Additionally, presented product is implemented with additional notification system and mirror-screen technology.

2.1 *Inertial Measurement Unit*

Inertial navigation sensors are necessary part of presented product. They expand possibilities of the smartphone. Each sensor is equipped with gyroscope, accelerometer and magnetometer. All of them are working in three perpendicular axes, shown in Fig. 1.

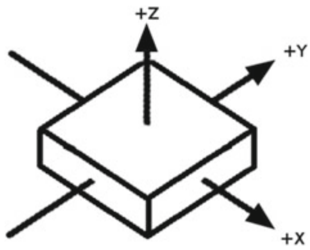


Fig. 1 Axes layout of a single IMU module

2.2 IMU Based Motion Capture Suits

Motion capture (mocap) systems available on the market differ from each other in price, quality of obtained measurements (which also depend on algorithms used to process the signal), quantity of sensors and method of their connection. Most often the set of sensors is attached to a cloth suit, creating a ready-to-wear motion capture solution.

In contrast to commonly used wired protocols (standard RS-232, RS-485), presented prototype uses wireless solutions based on Bluetooth and Wi-Fi technologies.

2.3 IMUs Data Output

Raw signals given by sensors are: angular acceleration, linear acceleration and magnetic field magnitude. Values are measured individually for each of three perpendicular axes.

Most of market available motion capture systems based on inertial navigation sensors give orientation of the individual module defined as rotation unit quaternion.

Quaternions are number system that extends the complex numbers, presented in computer science as four floating point numbers. A unit quaternion is a quaternion of norm one. Such quaternions can be identified with rotations in three dimensional space, which in combination with the specific properties of their algebra makes them a convenient tool.

A quaternion is defined by the following formula:

$$q = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} \tag{1}$$

where a, b, c are real numbers and $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ are fundamental quaternion units ($\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{ijk} = -1$). A quaternion can be treated as a pair of angle and the vector around which the rotation is made. Assuming an angle of rotation θ, and a norm one vector $\mathbf{u}$ representing the rotation axis, the scalar part of quaternion $a = \cos(\frac{\theta}{2})$ and $(b, c, d) = \sin(\frac{\theta}{2}) \cdot \mathbf{u}$ is the vector imaginary part [3].

3 Case Study—System Prototype

3.1 Basic Functionalities

The main objective of the system is to allow the user to exercise at any time using wireless IMU sensors connected to the smartphone. After initial interpretation, the measurements made by the sensors are sent to the smartphone in the form of orientation quaternions. These quaternions are then used to overwrite the rotation values of the corresponding parts of the avatar. Two animated avatars are displayed on the phone screen, one of which presents the exercise performed correctly and the other one is animated by data provided by inertial measurement units.

Before starting the exercise, user can familiarize himself with its course by observing an animated avatar driven by a reference motion capture recording. The angle and position of the visualization camera can be freely changed, allowing the user to observe exercise from any perspective. Playback of the animation allow pausing, rewinding, speeding up and down, and is seamlessly looped, maintaining the interface's transparency as in Fig. 2.

Positions of predetermined control points placed on the hands, knees, elbows and feet of avatars are compared, resulting in an array of error vectors. Based on the vector magnitude, indicator of exercise quality located on control points of avatar goes from green to red (where green stands for small magnitude, and red for large). Indicators are represented by primitive shape e.g. sphere or cube as in Fig. 3. After exceeding the configurable error threshold, an audio message with the suggested posture correction is played.

Selected characteristics of the course of the exercise performed by the user should be able to be saved for later interpretation by a physiotherapist.

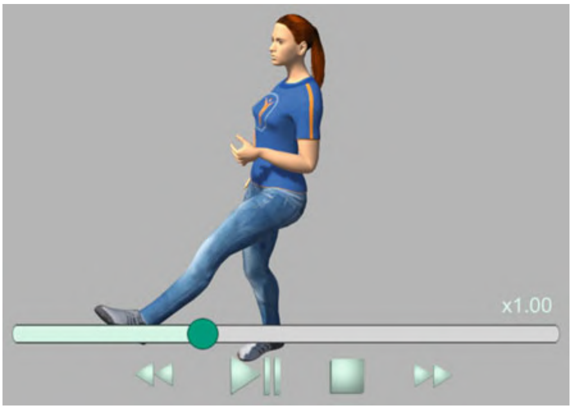
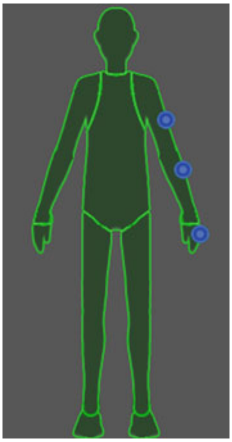


Fig. 2 Prototype appearance of the exercise preview screen

Fig. 3 Prototype visualization of error indicators during execution

Fig. 4 Location of IMU sensors during arm exercises



3.2 Hardware Requirements

Discussed prototype assumes the use of smartphone with Android 5 or later version, connected with wireless motion capture system based on inertial navigation sensors. Quantity of sensors is not strictly determined but it depends on type of the exercise. Number of sensors required for activity involving whole body oscillates near seventeen while hand to arm exercise requires only three sensors placed as shown in Fig. 4. Market available wireless IMU sizes are similar to a box of matches what qualifies them as fully mobile devices.

3.3 Reference Recordings

Application enables user to get acquainted with exercise guidelines by watching the display of animated avatar. Accuracy of exercise is determined using the difference of predetermined control points distances between reference and controlled avatar (controlled by user's moves). It is possible to derive error vector for every pair of reference and control points according to formula (2). That fully explains why preparation of reference recordings is crucial for every provided exercise. Recordings are prepared using highly accurate motion capture system, under the supervision of professional physiotherapist. To improve aesthetic values of recordings, animations are processed and adjusted for looping. Properly prepared files are exported to popular in Unity environment format *fbx*.

$$ERROR(A, B) = \begin{bmatrix} B_x \\ B_y \\ B_z \end{bmatrix} - \begin{bmatrix} A_x \\ A_y \\ A_z \end{bmatrix} \tag{2}$$

where

$$A, B \text{ - compared position vectors}$$

3.4 Layer Divided System

To fully explain the idea of prototype it is necessary to separate it into three following layers which are presented in Fig. 5:

- motion capture layer,
- user device layer,
- application layer.

The separation is made with prototype function determinant. It allows to retain full flexibility during final system design.

3.5 Motion Capture Layer

Prototype involves wireless system with adequate number of IMU sensors to a chosen exercise. Each sensor, being part of the mocap suit, returns its current unit rotation quaternion. Data are transmitted further to the smartphone via Bluetooth or WiFi.

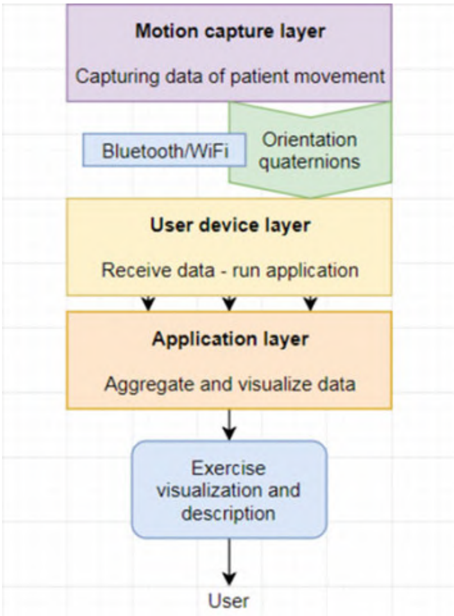


Fig. 5 Layers of the system

3.6 User Device Layer

According to the assumption that inertial navigation modules are wireless, smartphone aggregates data using Bluetooth or Wi-Fi technologies. The use of smartphone takes place through touchscreen, whose size and resolution have an influence on application readability. Exercise progress and application use are supported by voice notifications, consequently the speaker or headphones output are necessary. Smartphone device is recommended for prototype development because of it's mobility however there are no set up limits for device selection as long as it fulfills hardware requirements.

There is a possibility to show application on TV screen by using mirror screen technology preceded by launching the appropriate version of application on computer.

3.7 Application Layer

Prototype of rehabilitation application was prepared by using integrated Unity game engine environment. It allows relatively hassle-free change of target platform and in addition it delivers all needed equipment for visualisation. Communication interface is completely separated from application logic to provide problem-free integration with different motion capture technologies, independently from communication pro-

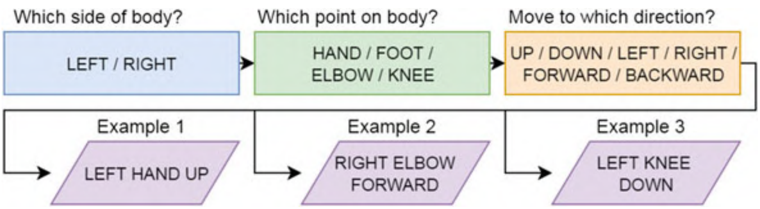


Fig. 6 Voice notifications assembling scheme

tocol. Environment chosen by authors enables seamless rendering of two animated avatars at once on smartphone screen, audio support and easy access to C# libraries.

3.8 Voice Notifications

To achieve versatility of evaluation criterion and voice notifications, the authors decided to examine the deviation between control points discussed in Sect. 3.3. In the program, deviation is implemented as a vector, what results in getting additional information about direction of the error. By processing such vector, it is easy to determine how to move part of body to which the control point was attached. Based on the name of the control point and its error vector full messages are created. The message preparation procedure starts with determining the body side, checks which checkpoint applies, then calculates the direction of movement improvement (see: Fig. 6). Obtained content of message can be read by text-to-speech technology or by the combination of sound samples with recorded speeches.

Not all messages are created based on control points—every exercise got it's exclusive set of messages proposed by the physiotherapist in the process of planning the exercise. Example exercise guidelines are presented in Table 1. There are also messages not connected directly with the course of exercise, e.g. information about connection or discharge of motion capture costume.

3.9 Excercises Files

During the implementation, it was assured that new exercises can be added and edited effortlessly. Each exercise corresponds to a single structure built according to the diagram in Fig. 7. A single exercise file contains information about the name of the exercise, the location of the IMU sensors on the patient's body, a reference recording of the exercise run, prepared as described in the Sect. 3.3, and its short text description. The set of files embedded in the application is called excercise pack.

Table 1 Example guidelines of exercise

Exercise phase	Phase description	Initial messages	Correction commands
(Starting position) Stand up straight	Feet slightly apart, knees straight, upper limbs on the hips, head straight	'Stand up straight' 'Upper limbs on the hips'	N/A
(Motion) Alternate bending and straightening of the knee joints	Lower limbs alternately bend in the knee and hip joints, lifting the knees to the hip	'Bend your knees to your hips' 'Alternately, right and left limb'	'Knees higher!' 'Hands on hips!' 'Straighten the spine!' 'Head straight!'
(Final position) Stand up straight	N/A	N/A	N/A

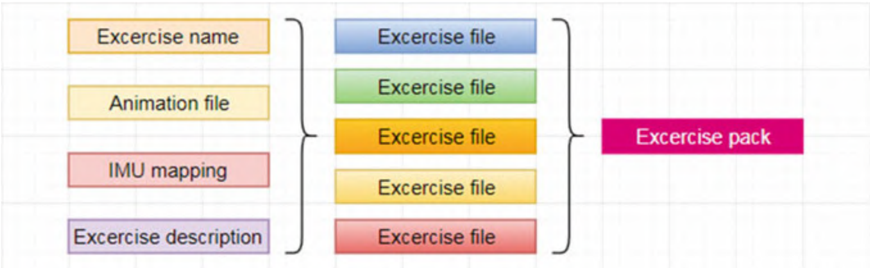


Fig. 7 Exercise files scheme

Depending on the scope of patient rehabilitation, the application will be equipped with an excercise pack corresponding to her/his disease. Creating sets of exercises is unrestricted.

3.10 Data Aggregation

Data flowing from the inertial navigation sensors to the main application can be aggregated and reviewed later by the patient or physiotherapist. It is possible to save data concerning not only the time of the exercise, but the number of repetitions, type of exercises and the exact course of each movement. Such information provided to a physiotherapist can help in the verification of disorders in the field of motor skills of a given patient.

3.11 Ubiquitous Rehabilitation

One of the assumptions of the application is to enable the patient to rehabilitate ubiquitously. This is achieved not only by the small size of the necessary devices, but also through the notification system. The mobile application installed on the patient's phone is enabled to send notifications that may remind the user to perform the exercise at set time intervals. Such a system of reminders can significantly improve the regularity of exercises.

3.12 Tests

Motorola Moto G3 was used during the tests. Custom system for capturing motion was used, based on three inertial navigation modules communicating over the air via the Wi-Fi standard.

The IMU sensors were refreshing their orientation quaternions with a frequency of not less than 60 Hz, what allowed them to detect even sudden movements. The issues of time synchronization between module measurements, processing raw data into an orientation quaternion, and transmission of the combined measurements of all the suit sensors into single packets has been ensured on the part of the used motion capture system.

The load test of the application was performed, in which set of 10, 100, 200 and 400 modules were simulated in succession. Considering the fact that the most popular suits for motion capture use less than 40 modules, test results show that Application Layer does not limit the speed of the system. The main factor determining the delays in the prototype is the system used for motion capture and the characteristics of its data transmission. The delay between the user's movement and its visualization on the three-dimensional avatar is shorter than two refreshment cycles of the orientation quaternion. The performance test screen and results are shown in Figs. 8 and 9.

The tests did not focus on the specifications of used motion capture solution due to the fact that the application is universal and can be combined with any Wi-Fi suit returning an orientation quaternion.

Starting the early phase of the implementation, a test environment was created. It simulates the operation of the IMU-based system which allows the verification of the application's assumptions without connecting to the actual motion capture device. Application tests, during which the operation of the implemented functionalities was checked, were successful. The measurement accuracy of the motion capture suit as well as its transmission speed have reached the values predicted for this system. A basic set of 10 exercises was created covering the range of the entire human body.

Implementation and subsequent tests were a valuable help for planning the further development described in Sect. 4.

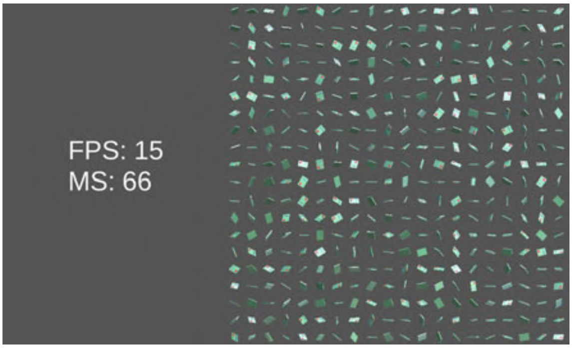


Fig. 8 Performance test executed on smartphone

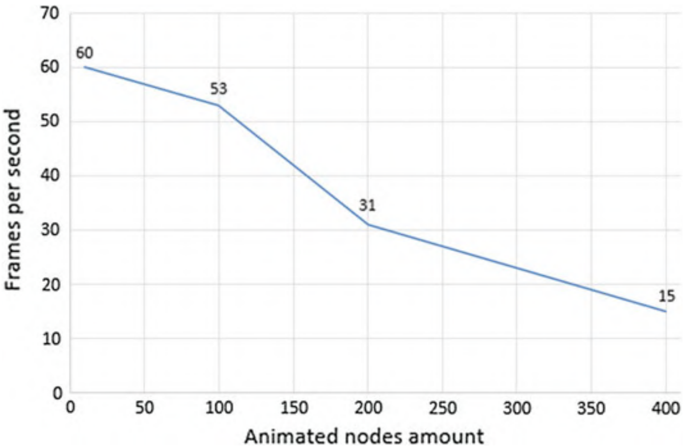


Fig. 9 The effect of the animated modules number on screen refresh rate

4 Further Development

Currently, the application is at the stage of an early prototype. Testing phase described in Sect. 3.12 allowed to verify the assumptions and to choose the appropriate direction for further development of the prototype. It is planned to extend the set of exercises with additional packages, not only related to rehabilitation processes, but also to recreational purposes. The aesthetic improvements of the interface and avatars will be introduced. The application will be supported by detailed documentation. System integration tests will be also carried out with various motion capture devices available on the market. After performing described activities and removing any detected errors and conflicts in the software, research will be conducted on a wider group of recipients.

5 Conclusions

The paper presents a prototype of a standalone solution allowing for independent rehabilitation using fully portable devices, such as smartphone and wireless IMU sensors.

Data transmitted from inertial navigation modules are received by smartphone, then interpreted and visualized in the form of an animated avatar. The high portability of the solution allows exercises to be performed anywhere, at any time, enabling the rehabilitation to be ubiquitous.

Presented system helps the user get familiar with the movements performed during the exercise by previewing them on the animated avatar. The proposed solution provided voice notifications that facilitate the operation of the software and the course of rehabilitation. Using the functionality of later versions of Android, it is possible to display the application's view on the TV screen via the mirror-screen function.

The built-in notification system supports regularity of the exercises. The data aggregated by the smartphone related to the patient's movement can be reviewed and interpreted by the specialist supervising patient's health.

Due to isolation of the visualization logic and the communication interface between the phone and the motion capture system, this solution is fully universal and can be adapted to various types of inertial navigation devices.

While designing the solution, usage of Android system was assumed, but the chosen software development environment allows to transfer the application to other popular platforms such as Windows and iOS.

Acknowledgements Data used in this project were obtained from the Centre for Research and Development of the Polish-Japanese Academy of Information Technology (PJAIT) (http://bytom.pja.edu.pl/).

This work has been supported by the National Centre for Research and Development "Innovative IT system to support training in alpine skiing and biathlon, with the functions of multimodal motion data acquisition, visualization and advanced analysis using machine learning techniques. Snowcookie PRO".

This work has been supported by Ministry of Science and Higher Education, PhD implementation program.

References

1. Uttarwar, P., Mishra, D.: Development of a kinect-based physical rehabilitation system. In: 2015 Third International Conference on Image Information Processing (ICIIP), pp. 387–392. Waknaghat (2015)
2. Deponti, D., Maggiorini, D., Palazzi, C.E.: DroidGlove: an android-based application for wrist rehabilitation. In: 2009 International Conference on Ultra Modern Telecommunications & Workshops, pp. 1–7. St. Petersburg (2009)
3. Świtoński, A., Josiński, H., Wojciechowski, K.: Dynamic time warping in classification and selection of motion capture data. Multidimens. Syst. Signal Process. (2018) (in Press)

4. Chutatape, O., Naonueng, K., Phoophuangpairoj, R.: Detection of improper postures leading to dislocation of hip prosthesis by a smartphone. In: 2017 14th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), pp. 415–418. Phuket (2017)
5. Kontadakis, G., Chasiouras, D., Proimaki, D., Mania, K.: Gamified 3D orthopaedic rehabilitation using low cost and portable inertial sensors. In: 2017 9th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games), pp. 165–168. Athens (2017)
6. Pinto, J.F., Carvalho, H.R., Chambel, G.R.R., Ramiro, J., Goncalves, A.: Adaptive gameplay and difficulty adjustment in a gamified upper-limb rehabilitation. In: 2018 IEEE 6th International Conference on Serious Games and Applications for Health (SeGAH), pp. 1–8. Vienna (2018)
7. Frederix, I., Sankaran, S., Coninx, K., Dendale, P.: MobileHeart, a mobile smartphone-based application that supports and monitors coronary artery disease patients during rehabilitation. In: 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) pp. 513–516. Orlando, FL (2016)
8. Medvedev, O., Marshall, A., Antonov, A.: User-friendly interface for the smartphone-based self management of pulmonary rehabilitation. In: 2008 International Conference on BioMedical Engineering and Informatics, pp. 673–676. Sanya (2008)
9. Maradani, B., Levkowitz, H.: The role of visualization in tele-rehabilitation: a case study. In: 2017 7th International Conference on Cloud Computing, Data Science & Engineering—Confluence, pp. 643–648. Noida (2017)
10. Dallaire-Cote, M., Charbonneau, P., Cote, S.S.P., Aissaoui, R., Labbe, D.R.: Animated self-avatars for motor rehabilitation applications that are biomechanically accurate, low-latency and easy to use. In: IEEE Virtual Reality (VR), pp. 167–168. Greenville, SC (2016)
11. Moya, S., Grau, S., Tost, D., Campeny, R., Ruiz, M.: Animation of 3D avatars for rehabilitation of the upper limbs. In: 2011 Third International Conference on Games and Virtual Worlds for Serious Applications, pp. 168–171. Athens (2011)
12. Gunaydin, T., Arslan, R.B., Birik, B., Cirak, Y., Durustkan Elbasi, N.: A surface electromyography based serious game for increasing patient participation to physiotherapy and rehabilitation treatment following anterior cruciate and medial collateral ligaments operations. In: 2018 IEEE 31st International Symposium on Computer-Based Medical Systems (CBMS)

Decision Making Based on IoT Data Collection for Precision Agriculture

Christine Dewi and Rung-Ching Chen

Abstract Internet of Things (IoT) is a shared network of things which can associate with each other through the internet connection. IoT plays a vital role in the agriculture industry which can feed 9.0 billion people by 2050. Precision Agriculture provides a novel solution using a systems approach for today's agricultural issues such as the need to balance productivity with environmental concerns. This paper proposes IoT for local information data collection on Precision Agriculture and uses some case study for examples. In this paper, we aim that the internet of things can be used to collect local information data on precision agriculture. The farmer could get the real time data for monitoring his field. Moreover, by using this technology, farmers can effectively use the information to achieve higher yields and therefore earn higher profits.

Keywords Internet of things · Local information · Data collection · Precision agriculture

1 Introduction

As of now the total population of the world is presumed to be approximately 7 billion and is projected to increase by 2.0 billion in just 40 years, reaching 9.0 billion by 2050 [1]. This global resident growth affects the problem of food shortfall, and it becomes a world-scale problem. As the world's population grows, farmers will need to supply more food. Moreover, there are some innovations in the agricultural field, for example, the concept of Precision Agriculture. Precision Agriculture provides a novel fulfillment using a systems approach for today's agricultural issues such as

C. Dewi · R.-C. Chen (✉)
Department of Information Management, Chaoyang University of Technology, Taichung, Taiwan, Republic of China
e-mail: crching@cyut.edu.tw

C. Dewi
e-mail: s10714904@cyut.edu.tw

M. Huk et al. (eds.), *Intelligent Information and Database Systems: Recent Developments*, Studies in Computational Intelligence 830,
https://doi.org/10.1007/978-3-030-14132-5_3

the need to balance productivity with environmental concerns. It is based on the implementation of foremost information technologies. It includes describing and modeling variation in soils and plant species and integrating agricultural practices to meet site-specific requirements.

Furthermore, Precision Agriculture is the concept of using new technologies and collected field information. The main activities of Precision Agriculture are data collection, processing and variable rate applications of inputs. In addition, Internet of Things (IoT) technology is currently shaping different aspects of human life. Precision agriculture is one of the paradigms which can use the IoT advantages to optimize the production efficiency across the agriculture fields, optimize the quality of the crops, and minimize the negative environmental impact [2]. The use of Wireless Sensor Networks (WSNs) in precision agriculture upgrades the overall efficiency, productivity, and profitability of the agricultural system. WSNs are key components the IoT in which different pieces of information gathered from almost anywhere and anything in the world is accessible through the Internet [2]. Of the various advantages that IoT brings to the table, its ability to innovate the current scenario of farming methods is absolutely ground-breaking. Mostly, we come across ideas that suggest a wireless sensor network that collects data from the various sensors present in the field and sends the data to the main central server. This method focuses on studying the environmental factors to improve crop yield [3].

This paper proposes IoT for local information data collection on Precision Agriculture. This paper aims to implement IoT technology to get the data for monitoring systems such as crop and soil monitoring, weather and air monitoring and working machine monitoring. The remainder of the paper is organized as follows. In Sect. 2, we review the related work and the proposed local information collection. The case study of IoT architecture is presented in Sect. 3. The paper is concluded in Sect. 4.

2 Related Work

There are many types of research about the use of Internet of Things on Precision Agriculture. Prasad et al. [4] propose an approach combining the advantages of the significant characteristics of emerging technologies such as the Internet of Things (IoT) and Web Service to construct an efficient approach to handle the enormous data involved in agrarian output. The method uses the combination of IoT and cloud computing that promotes the fast development of agricultural modernization and helps to realize the smart solution for agriculture and efficiently solve the issues related to farmers. This research use connectivity using existing 3G, 4G networks using the available hardware is one significant advantage for Smart agriculture. Internet of things paradigms are used as resources to facilitate decision making. IoT architecture, operating rules, and smart processes are implemented using a distributed model based on edge and fog computing paradigms. A communication architecture is proposed using these technologies. The aim is to help farmers to develop smart systems both, in current and new facilities [5]. The important things to do in this paper describe how to collect the local information data use the IoT before analyzing the other step.

The Internet of Things (IoT) is one network which contains physical objects that are embedded with the electronic devices, software, connectivity and sensors to achieve a higher value and provide some services regarding the exchange of details with the product manufacturer, operator [4]. IoT is developed using architectures based on layers capable of connecting a huge number of devices with each other and with the established services. The basic model of IoT has a three layer architecture which is of Perception, Network and Application Layers [5]. Protocols are important depends on several facts such as environmental conditions, network characteristics, the total of data to be transferred, quality of service requests and security levels [6]. Furthermore, in this work different protocols such as Hypertext Transfer Protocol (HTTP), Message Queuing Telemetry Transport (MQTT), Bluetooth, and Wifi can be used to develop proposed architecture. IoT technologies are proposed in PA scenarios. Moreover, in [7] this model is described as a solution in precision farming. IoT Smart farming application include farm parameters tracking, monitoring, field observation, and storage monitoring. The work Internet of Things Platform for Smart Farming presents a platform based on IoT technologies that can automate the collection of environmental, soil, fertilization, and irrigation data; automatically correlate such data and filter-out invalid data from the perspective of assessing crop performance; and compute crop forecasts and personalized crop recommendations for any particular farm [6].

2.1 Local Information Data Collection

Local information system describes in this paper can see in Fig. 1. The data from the local system divided by three part, first crop and soil monitoring, weather and air monitoring, and last working machine monitoring.

Data are facts and statistics collected together for reference or analysis. Data collection is the most important in precision agriculture because the farmer can manage

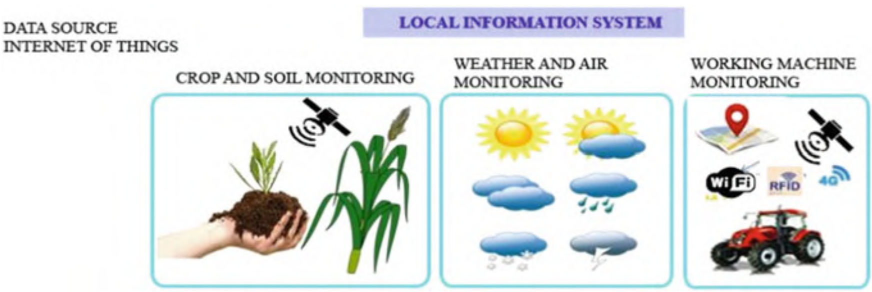


Fig. 1 Data source internet of things on local information system

their crop real time and improve their crop productivity. There many research about precision agriculture in a different crop. In [8] explain and give an example about yield monitoring over time creates a unique GIS database that assists farmers to easily identify yield variability within a field, to make better variable-rate decisions, and create a history of spatial field data. Yield monitoring has become a common practice in traditional grain crops and corn-soybean rotation systems. This technology is being researched and commercialized for other crops such as potato, onion, sugar beet, tomato, hay, citrus, grape, and sugarcane. There are various elements of precision farming that play their necessary role in cropping systems. Yield monitoring is being used to measure the yield variability of corn, soybean, potato, tomato, onion, sugar beet, hay, orange, grape, sugarcane and has become the most widely used component of precision farming. Besides yield variability, crop variety comparisons, yield damage reports, and field efficiency are being assessed with the use of yield monitoring systems.

Soil preparation is the first step before growing a crop. The ultimate objective is to produce a firm and weed-free seedbed for rapid germination and emergence of the crop. One of the most important tasks in soil preparation is tilling (or ploughing): turning the soil and loosening it. Besides, soil preparation is one of the most energy-consuming parts of agriculture, requiring significant inputs of fuel and time. Depending on the field's location, it may also increase the risk of soil erosion. Nowadays, precision farming equipment exists that helps farmers to use considerably less fuel and time in soil preparation by improving the accuracy, efficiency, and sustainability of the process.

According to Pavithra [9] the distinctive sensors which are utilized in the research are the Temperature sensor, Humidity sensor, Soil sensor, Water-level sensor, and Light Dependent Resistor can be seen on the Fig. 2.

Sensed data inputs produce several data. First, Temperature control is considering high radiation, while the temperature is evolving, and a few yields may get harmed, so it needs some ventilation strategy to control the temperature. Depending on the agricultural product being grown temperature can affect growth, germination, sprouting, flowering, and fruit development. Next, Water vapor is the fundamental issue that is influencing the development of products. Due to high humidity, odds of the ailment are expanding. Mugginess may affect hydria push, shutting the stomata and hence it might let down the procedure of photosynthesis which relies upon the osmosis. Moreover, soil water likewise influences the yield development. The best possible water systems and preparations of the product are differed according to the sort, age, stage, and atmosphere. Some key parameters for water level are pH esteem, dampness contains, electric conductivity and the temp of dirt. Furthermore, the Light dependent resistor is a photograph conductive sensor [9]. Based on Gaikwad and Galande [10], a wireless sensor network (WSN) is an infrastructure comprised of sensing, computing and communication elements that allow the administrator to monitor and control of the specified parameters in the network. A typical application of WSN includes data collection, monitoring, surveillance and medical telemedicine. It is also used in the irrigation system, Greenhouses for monitoring and controlling parameters like

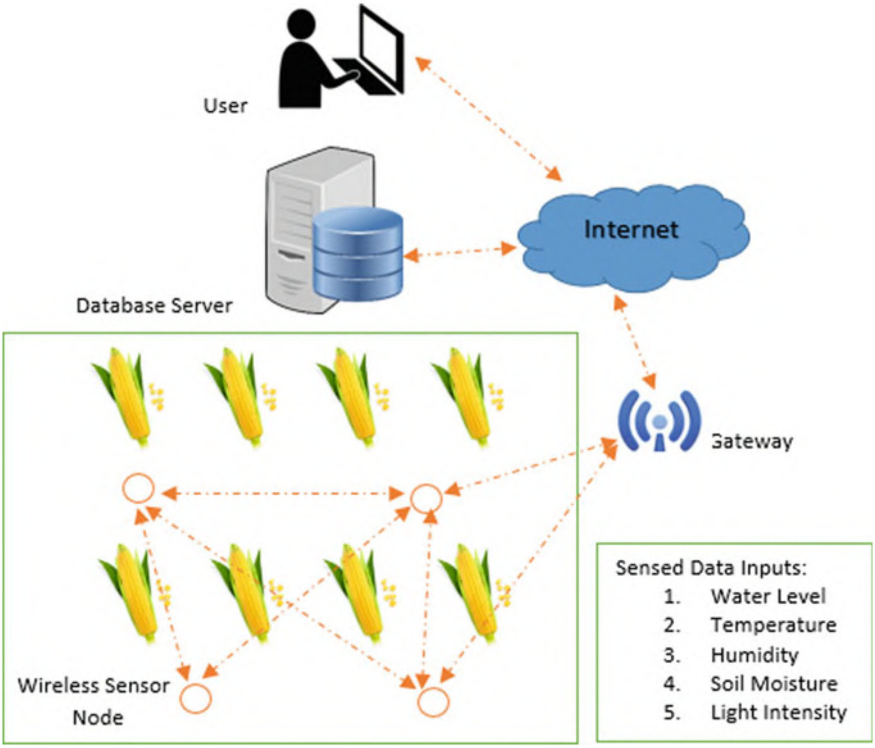


Fig. 2 Local information system architecture

water flow, temp, humidity, moisture, etc. Based on the previous research, the type of sensor for local information data collection could be seen in Table 1.

There were many type of sensors that can be used to get local information data like temperature sensor (LM35, DHT11, SHT75, SHT-3x DIS), humidity sensor (P-Hs-220, DHT11, HH10D, CMOS chip, SHT-3x DIS), Soil sensor (SN-M114, 10HS, DHT11), Water-level sensor (pH measurement sensor), and Light Dependent Resistor. This research proposed a system model that could be seen in Fig. 3.

Figure 3 shows the system model on local data information to collect data from the various crop (Potato, Onion, Sugar Beet, Tomato, Hay, Citrus, Grape, and Sugarcane, etc.). First, the process starts with sensors detection such as water level, temperature, humidity, soil moisture, and light. The data will collect, process, and save on database. The next step, the data will be processed and create a decision model (for example Decision tree and Markov). Furthermore, Decision tree is an efficient tool in machine learning for classification problems. The decision tree is based on a hierarchical decision scheme or a tree like structure. Each node of the decision tree structure makes a binary decision that separates either one class or some of the classes from the remaining classes. The processing is generally carried out by moving down the tree until the leaf node is reached. This is known as a top–down approach. Moreover,

Table 1 Type of Sensors based on previous research [10–16]

No.	Research	Soil moisture	Temperature	Humidity	Water level	Light intensity measurement
1.	Measurement of NPK, temperature, moisture, humidity using WSN [10]	Soil moisture sensor	Temperature sensor LM35 series	P-Hs-220 humidity sensor	pH measurement sensor	Light dependent resistor
2.	Field monitoring using IoT in agriculture [11]	Soil moisture sensor DHT11	Temperature sensor DHT11	Soil moisture sensor DHT11	–	–
3.	IoT based crop-field monitoring and irrigation automation [12]	Soil moisture Sensor	Temperature sensor DHT11	Temperature sensor DHT11	–	Light dependent resistor
4.	Wireless monitoring of soil moisture, temperature and humidity using Zigbee in agriculture [13]	Soil moisture sensor	Temperature sensor LM35 series	Humidity sensor	–	–
5.	IoT based smart soil monitoring system for agricultural production [14]	pH sensor	Temperature sensor LM35 series	HH10D relative humidity sensor	pH sensor	–
6.	Development of wireless sensor network for Harumanis Mango Orchard's temperature, humidity and soil moisture monitoring [15]	SN-M114 soil moisture sensor	Temperature sensor SHT75	CMOS chip humidity sensor	–	–
7.	Station for soil humidity, temperature and air humidity measurement with SMS forwarding of measured data [16]	10HS soil sensor	SHT-3x DIS temperature sensor	SHT-3x DIS humidity sensor		

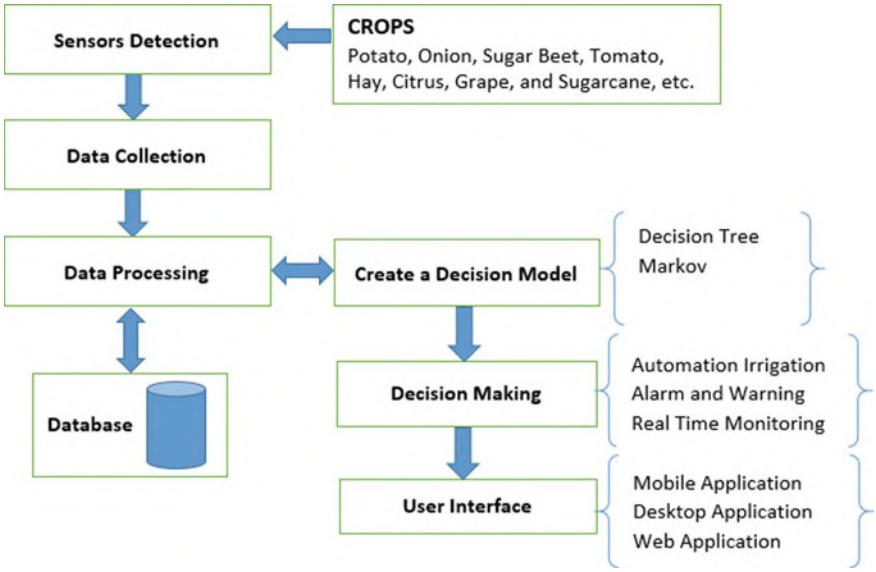


Fig. 3 System model

this analysis will help the farmer to monitor his crop real time. The system could give warning and alarm if something happened in the field as the change of temperature and humidity, so the farmer could make an action to solve this problem quickly, for instance, automation irrigation. The decision model analysis can help the farmers maximize their profit margin by providing predictions on crops that will give the maximum output in a particular area and in the future.

3 Case Study

Chavan and Karande [12] focuses on field monitoring using IoT devices which would provide real-time soil moisture, humidity, and temperature of the field to the farmers. An Arduino Microcontroller board with soil, temperature and humidity sensors is used to collect the data from the filed on the fly from a remote field. The sensors are deployed in the wheat crop for data capturing. The results are carried out and analyzed in the form of temperature, moisture and humidity from the wheat crop field for the invocating alarm to the farmer when required.

Figure 4 shows that as the temperature raises the humidity decreases alarmingly. Near 110 days the humidity falls to around 35. This again can be a trigger an alarm to the farmer to water the fields. Since the system can be on auto mode, the pump could start on its own. Moreover, the other case study describes wireless monitoring of soil moisture, temperature and humidity using Zigbee in Agriculture [13]. This

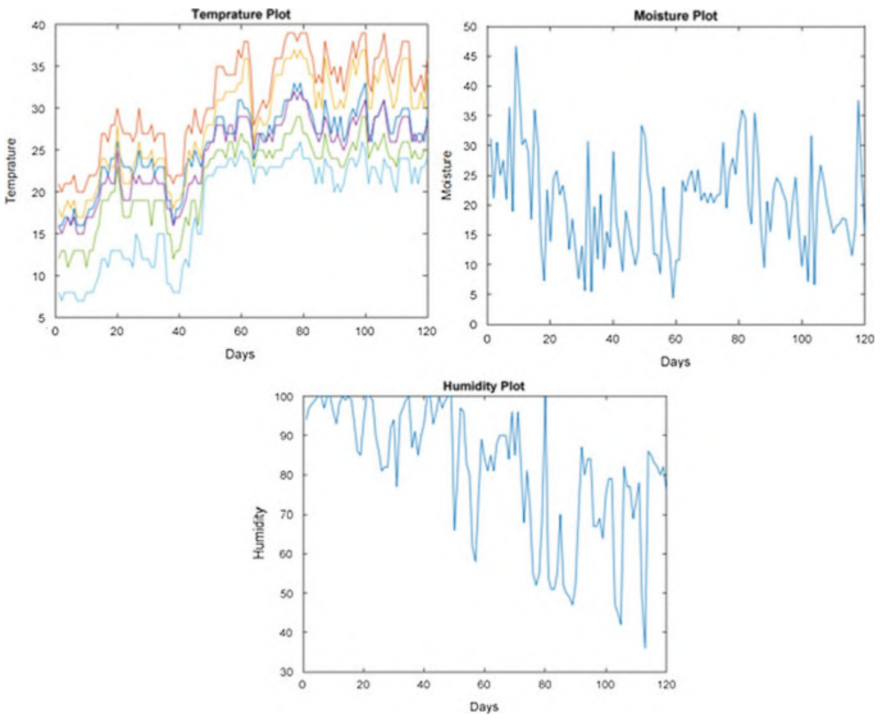


Fig. 4 Temperature, humidity and moisture plot analysis [12]

research concludes Zigbee-based agriculture monitoring system serves as a reliable and efficient system for monitoring agricultural parameters. The corrective action can be taken. Wireless monitoring of field not only allows the user to reduce the human power, but it also allows the user to see accurate changes in it. It is cheaper and consumes less power. Nirmal Kumar et al. [17], presented a real-time paddy crop field monitoring using Zigbee Network. This paper proposed an idea about monitoring the crop field area without human interaction. Besides, this research use four Zigbee nodes to form a Zigbee network. Temperature sensor and pressure sensor were connected to two Zigbee devices via microcontroller and humidity sensor was connected to another Zigbee via a microcontroller. The result of the experiment can be seen in Table 2 and Fig. 3.

The graph on Fig. 5 describes the temperature and humidity readings when implemented in real-time environment. The result shows that the temperature and humidity is a varying quantity in the paddy crop field area. A farmer can use this data to make decision support in the future to improve their productivity. For examples, Veenadhari et al. [18] focused on decision tree to evaluate the effect of climatic factors on the soybean crop yield. The study implement the potential use of this process for extracting useful information from existing secondary data of climatic factors. The decision trees suggested there exists a correlation between climatic factors and soy-

Table 2 Temperature and humidity reading [17]

No.	Time(s)	Temperature (°C)	Humidity (%)
1.	11.15	26.2	65
2.	11.16	26.1	64
3.	11.17	27.4	71
4.	11.20	26.9	59
5.	11.21	25.9	69

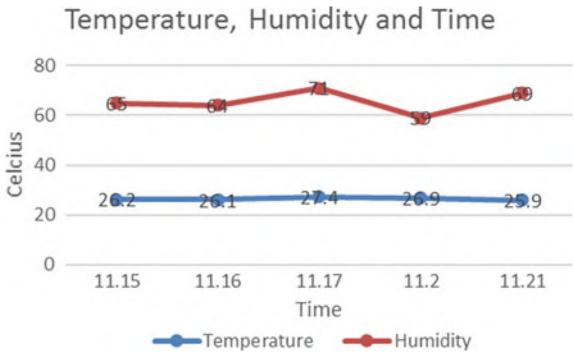


Fig. 5 Temperature and humidity analysis

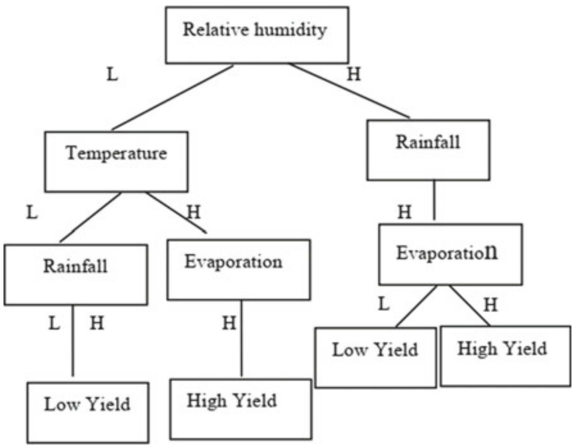


Fig. 6 Decision tree for influence of climatic factors on soybean yield [18]

bean crop productivity and these variables influence on the soybean crop productivity were confirmed from the rule accuracy. The decision tree analysis indicated that the productivity of soybean crop was mostly influenced by Relative humidity followed by temperature and rainfall. Figure 6 shows decision tree for influence of climatic factors on soybean yield.

Based on Shakoor et al. [19], in most of the cases, Decision Tree Learning-ID3 algorithm gives a less value for percentage error than the KKNR algorithm without omitting the outliers of the dataset. The research suggests area based beneficial crop

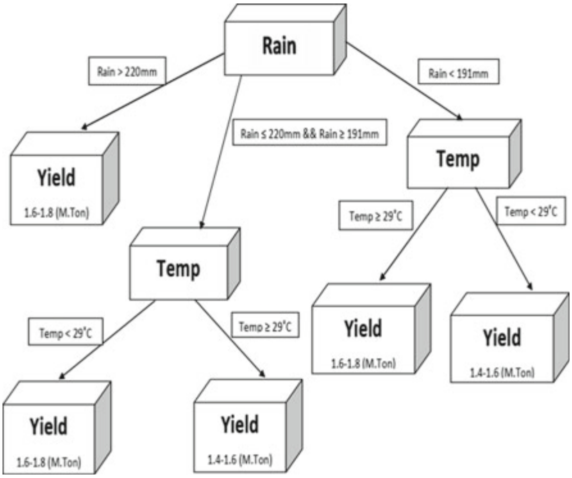


Fig. 7 Example of decision tree implementation [19]

rank before the cultivation process. The prediction is based on analyzing a static set of data using Supervised Machine Learning techniques (Decision Tree Learning-ID3 and K Nearest Neighbors Regression algorithms). ID3 uses the decision tree table which consists of the ranges of the rainfall, temperature and yield data. The goal of this research is to help the farmers maximize their profit margin by providing predictions on crops that will give the maximum output in a particular area. The dataset contains of details on crops' yield per hectare, average of minimum and maximum temperature, rainfall, year range, and region. Analyzing these data, the algorithms give the result which predicts the preferred profitable output. For the accuracy of the prediction, data of past twelve years is being used by the algorithms for the purpose of learning and result analysis. Example of decision tree implementation could be seen on Fig. 7.

In [20], a WSN-based outlier detection and irrigation management system is developed to assist farmers in handling crop irrigation schedules. The proposed system collects environmental and soil-related data through sensor nodes that are examined by the outlier detection module. The proposed noise detection algorithm is implemented to improve the system's accuracy. Afterward, DSS further examines it to identify water deficient sites in agricultural fields. After identification, the alarming unit is activated, and text messages are sent over a local area network (LAN) to inform farmers. The other research use a wireless decision support system for the optimized management of the irrigation in agriculture has been presented. The properties of the WSAN technology have been exploited to acquire heterogeneous environmental parameters and to control the functioning of the irrigation system. The FL-based methodology has been designed and calibrated according to the indications of the farmers in order to mimic the human experience and to properly understand the status of the crop [21]. Monitoring based on IoT is a reliable and fast system which helps farmers in monitoring the fields effectively. This helps them to take corrective measure for the protection and better yield of the crop.

4 Conclusion and Future Work

In this paper, we aim that the internet of things can be used to collect local information data on precision agriculture. The farmer could get the real-time data for monitoring his field. The important local information data are water level, temperature, humidity, soil moisture, and light from the various crop (Potato, Onion, Sugar Beet, Tomato, Hay, Citrus, Grape, and Sugarcane, etc.). There are many types of sensors could be used to get the important data in the agriculture field in Table 1. The decision model analysis, decision tree can help the farmers maximize their profit margin by providing predictions on crops that will give the maximum output in a particular area and the algorithms give the result which predicts the preferred profitable output. In the future, the monitoring system could be connected and information exchanged via sensors is available to the farmer on the website and desktop application and mobile phones. Furthermore, a farmer can use this data to make intelligent decisions for planting, fertilizing, and harvesting crops. Finally, using this technology, farmers can effectively use the information to achieve higher yields and therefore earn higher profits.

Acknowledgements This paper is supported by Ministry of Science and Technology, Taiwan. The Nos are MOST-107-2221-E-324-018-MY2 and MOST-106-2218-E-324-002, Taiwan.

References

1. Srinivasulu, P., Venkat, R., Babu, M., Rajesh, K.: Cloud Service Oriented Architecture (CSoA) for agriculture through the Internet of Things (IoT) and big data. In: 2017 IEEE International Conference on Electrical, Instrumentation and Communication Engineering (ICEICE), IEEE, Karur, India (2017)
2. Khattab, A., Abdelgawad, A., Yelmarthi, K.: 2016 28th International Conference on Microelectronics (ICM), IEEE, Giza, Egypt, pp. 201–204 (2016)
3. Amandeep, Bhattacharjee, A., Das, P., Basu, D., Roy, S., Ghosh, S., Saha, S., Pain, S., Dey, S.: 2017 8th IEEE Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), IEEE, Vancouver, BC, Canada, pp. 278–280 (2017)
4. Prasad, K., Kumaresan, A., Nageshwaran, B., Kumaresan, A., Kotteshwaran, M.: IoT based smart agricultural solutions to farmer enhanced with WIFI technology. Int. J. Adv. Res. Basic Eng. Sci. Technol. (IJARBEST) **3**, 131–136 (2017)
5. Ferrández-Pastor, F., García-Chamizo, J., Nieto-Hidalgo, M., Mora-Martínez, J.: Precision agriculture design method using a distributed computing architecture on internet of things context. Sensors (2018)
6. Hedi, I., Špeh, I., Šarabok, A.: IoT network protocols comparison for IoT constrained networks. In: Proceedings of the 2017 40th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Opatija, Croatia, pp. 501–505 (2017)

7. Tech-Target: IoT as a Solution for Precision Farming (2017). Available online: http://internetofthingsagenda.techtarget.com/blog/IoT-Agenda/IoT-as-a-solution-for-precision-farming. Accessed on 4 Oct 2018
8. Koch, B., Khosla, R.: The role of precision agriculture in cropping systems. J. Crop Prod. **9**(1–2), 361–381 (2003)
9. Pavithra, G.: Intelligent monitoring device for agricultural greenhouse using IOT. J. Agric. Sci. Food Res. **9**, 220 (2018)
10. Gaikwad, S.V., Galande, S.G.: Measurement of NPK, temperature, moisture, humidity using WSN. Int. J. Eng. Res. Appl. **5**(8), (Part-3), 84–89 (2015)
11. Krishna Jha, R., Kumar, S., Joshi, K., Pandey, R.: Field monitoring using IoT in agriculture. In: 2017 International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT), IEEE, Kannur, India (2017)
12. Chavan, C.H., Karande, P.V.: Wireless monitoring of soil moisture, temperature & humidity using Zigbee in agriculture. Int. J. Eng. Trends Technol. (IJETT) **11**(10) (2014)
13. Rajalakshmi, P., Mahalakshmi, S.D.: IOT based crop-field monitoring and irrigation automation. In: 10th International conference on intelligent systems and control (ISCO), IEEE, Coimbatore, India (2016)
14. Ananthi, N., Divya, J., Divya, M., Janani, V.: IoT based smart soil monitoring system for agricultural production. In: IEEE technological innovations in ICT for agriculture and rural development (TIAR), IEEE, Chennai, India (2017)
15. Nooriman, W.H., Abdullah, A.H., Abdul Rahim, N., Kamarudin, K.: Development of wireless sensor network for Harumanis Mango Orchard's temperature, humidity, and soil moisture monitoring. In: IEEE Symposium on Computer Applications & Industrial Electronics (ISCAIE), IEEE, Penang, Malaysia, pp. 264–268 (2018)
16. Bodić, M., Vuković, P., Rajs, V., Vasiljević-Toskić, M., Bajić, J.: Station for soil humidity, temperature and air humidity measurement with SMS forwarding of measured data. In: 41st International Spring Seminar on Electronics Technology (ISSE), IEEE, Zlatibor, Serbia (2018)
17. Nirmal Kumar, K., Ranjith, P., Prabakaran, R.: Real time paddy crop field monitoring using Zigbee network. In: 2011 International Conference on Emerging Trends in Electrical and Computer Technology (ICETECT), Nagercoil, India, pp. 1136–1140 (2011)
18. Veenadhari, S., Mishra, B., Singh, C.D.: Soybean productivity modelling using decision tree algorithms. Int. J. Comput. Appl. **27**(7), 11–15 (2011)
19. Shakoor, T., Rahman, K., Nasrin Rayta, S., Chakrabarty, A.: Agricultural production output prediction using supervised machine learning techniques. In: 2017 1st International Conference on Next Generation Computing Applications (NextComp), IEEE, Mauritius (2017)
20. Khan, R., Ali, I., Zakarya, M., Ahmad, M., Imran, M., Shoaib, M.: Technology-assisted decision support system for efficient water utilization: a real-time testbed for irrigation using wireless sensor networks. IEEE **6**, 25686–25697 (2018)
21. Viani, F., Bertolli, M., Salucci, M., Polo, A.: Low-cost wireless monitoring and decision support for water saving in agriculture. IEEE Sens. J. **17**(13), 4299–4309 (2017)

Meeting Smart City Latency Demands with SDN

Filip Holik

Abstract Smart cities utilise a large number of heterogeneous devices with goal to improve all aspects of city operations. This often includes critical functionality, which brings strict requirements on network communication. One of the most important requirements is network latency. This paper analyses latency requirements in areas of smart cities and related domains. Analysed requirements are then practically tested in use case networks based on software-defined networking—a modern paradigm of network programmability, which is becoming used in various areas, including smart cities. The goal of the testing is to verify, if software-defined networking can achieve analysed requirements and can be therefore safely utilised as a driving technology of smart cities. Realistic results are achieved by testing in two different scenarios: in a small scale network composed from real hardware devices, and in a large scale emulated network corresponding to a typical smart city topology.

Keywords Connectivity requirements · Network latency · OpenFlow performance · SDN latency · Smart city network requirements

1 Introduction

The goal of a smart city is to "*improve outcomes of any aspect of city operations*" [1]. In order to achieve that, smart cities integrate several application domains, which have various network requirements including: maximum throughput, high availability, resiliency, reliability, redundancy, interoperability, flexible configurability, and security. While these requirements can be contradictory and are therefore often not required at the same time, there is an important requirement, which is always required by all domains—low latency.

Latency is defined as a time interval between sending a message from node *A* and receiving the message on node *B*—it is therefore measured one-way. A two-

F. Holik (✉)
University of Pardubice, Studentska 95, 532 10 Pardubice, Czech Republic
e-mail: filip.holik@upce.cz

M. Huk et al. (eds.), *Intelligent Information and Database Systems: Recent Developments*, Studies in Computational Intelligence 830,
https://doi.org/10.1007/978-3-030-14132-5_4

way latency, also called Round-Trip-Time (RTT), is then defined as a time interval between sending a message and receiving the reply on *A*—the reply is sent by *B*.

Certain aspects of smart city communication require latency to be below a specified threshold, otherwise the proper functionality cannot be ensured. Latency is therefore a critical parameter for functions requiring 100% reliability. Software-defined networking (SDN) is being used in the area, but unfortunately there is currently no exhaustive testing of SDN influence on latency in smart city scenarios. This paper analyses several key functions of SDN and their capability to meet the smart city latency requirements. The results of the testing determine, if SDN can be applied in this demanding area and if the software nature of SDN can cope with hardware-based traditional networks.

The rest of the paper is organised as follows: Sect. 2 presents smart city architecture composed from several domains, Sect. 3 analyses smart city latency requirements of these domains, and Sect. 4 verifies these requirements in several testing scenarios.

2 Smart City Architecture

Current smart cities are built independently without a clearly defined architecture. The only common principle is a backbone network connecting various smart city domains. The most used domains in smart cities are:

- Smart buildings—their main goals are to reduce consumption of resources, integrate security mechanisms, and provide intelligent environment monitoring and control. They are composed from sensors, an IP network, and software applications. The network typically uses Ethernet and Wi-Fi, and connects the building to an ISP or straight to the smart city network.
- Smart grid networks—integrate traditional power grid with an information network [2] to achieve clean, safe, secure, reliable, resilient, efficient, and sustainable system [7]. Their architecture can be decomposed into five main domains [5]: generation, transmission, distribution, distributed electrical resources (DER), and customer premises. From the smart city perspective, the most important domain is distribution. This domain contains Ethernet-based substation networks, which have the most strict network requirements—especially on low latency.
- Smart homes—they can integrate smart grid equipment for energy monitoring and management with various consumer devices (smart TVs), appliances (washing machines), and security devices (lock systems). The network architecture typically includes a single central device (combination of modem, router, switch, and access point), which provides wired and wireless connectivity.
- Smart mobility—it covers all V2X (vehicle to vehicle/infrastructure) communication and aims to increase quality of life and productivity. This can be achieved with optimised traffic control, predicted arrival times, and real-time traffic information. The network architecture uses various wireless transmission protocols (including cellular) and it also integrates wired networks.

The backbone network then has to comply with requirements of specific domains using one of the following approaches (or their combination):

1. Dedicated traffic links—each domain or application with different network requirements, has its own backbone physical connection. This approach offers the best performance, independent of other traffic types. On the other hand, the realisation is expensive and often not practically achievable.
2. Converged link—a single link is shared between several domains and applications. To logically separate their traffic, virtualization techniques can be used. These can include VLANs, QoS, or custom made SDN-based techniques.

3 Smart City Latency Requirements

This section analyses the latency requirements of specific smart city domains. The most important latency requirements are summarised in Table 4.

3.1 Smart Grid Latency Requirements

The most critical part of smart grid from the latency perspective is communication of IEDs (Intelligent Electronic Devices). To ensure proper grid functionality, IEDs often have to quickly react to an event (such as a lightning strike). A delayed reaction can cause grid damage, blackouts, and even potential deaths. According to [13], this critical communication requires latency under **3 ms** (4 ms according to [8]). Fortunately, this communication is bounded only to a specific substation network. The smart city network handles only the general IED communication, which requires latency under **10 ms** [13] (8 ms according to [8]). The standard [8] also covers substation internal monitoring and control communications, which require latency below **16 ms**.

The remaining smart grid communication has much less strict requirements, specifically: medium speed control (100 ms), slow speed control (500 ms), maintenance (1 s), audio and video transfer (1 s), and text transfer (2 s) [8, 13].

3.2 Smart Mobility Latency Requirements

Specified maximum latency in vehicular networks is within **50–100 ms** [10]. This depends on one of the following communication types:

- V2V—is a safety-related communication, which requires low latency and high reliability. The maximum latency must be under **100 ms** (**50 ms** for pre-crash sensing), with the minimum frequency of transmitted messages being 10 Hz. The communication includes: vehicle status warning (emergency electronic brake lights, abnormal condition), vehicle type warning (emergency vehicle, slow vehicle, motorcycle, vulnerable road user), traffic hazard warning (wrong way driving, stationary vehicle, traffic condition, signal violation, roadwork), and dynamic vehicle warning (overtaking vehicle, lane change assistance, pre-crash sensing, co-operative glare reduction).

- V2I/I2V—is mostly a non safety-related communication with not so strict requirements. This type has the maximum latency specified to **500 ms** (except for optimal speed advisory for traffic lights and intersection management—these require latency below **100 ms**) and minimum frequency of messages being 2 Hz. The communication includes: traffic management (speed limits, optimal speed advisory for traffic lights, intersection management, co-operative flexible lane change, electronic toll collection) and infotainment (point of interest notification, local electronic commerce, media download, map download and update).

To meet these demands, the most suitable and used protocol is 802.11p, which can comply with 1.5 ms latency requirement even in real-world scenarios with speeds of vehicles between 30 and 170 km/h [3]. For this communication, cellular networks (even LTE) are not suitable [6].

3.3 Nodes Latency Requirements

Most of the smart city nodes and sensors are battery-powered and use a wireless communication. To limit power consumption, several types of power-saving modes are utilised. Typically, devices are in a sleep mode and wake-up only when there is a need to transmit or receive data. Some communications such as LoRa, even have to comply with laws regulating the maximum "time on air" (1% for LoRa according to Europe regulation limitations). Different wireless protocols have various requirements on connection establishment times and node wake-up times. These requirements are a good indicator of maximum network latency. An SDN application should comply to these demands. This ensures that each node will be correctly registered by the network. A slight random delay should not present a major problem and nodes should still be able to join the network. Only a major delay or a packet loss could result in a total loss of connectivity between the node and the application.

The most used protocols for nodes communication in smart cities have the following time specifications: 6LoWPAN (**5 ms** [9]), WiMAX (**100 ms** [11]), and Zigbee (**30 ms** for connection establishment, **15 ms** for nodes wake-up [14]). Additionally, nodes can be connected via cellular networks. Depending on the generation, these networks have the following latency: GPRS (**150 ms**), UMTS (**100 ms**), LTE (**20 ms**), and 5G (expected to be **2 ms**) [12].

3.4 General Smart City Application Latency Requirements

The general applications, which demand low latency are: games streaming (**10 ms**), online games (**20 ms**), cloud applications (**50 ms**), VoIP and video conferencing (**150 ms**). They are used in smart homes and smart buildings such as: offices, government buildings, hospitals, schools, and public buildings (museums, theatres, etc).

3.5 Symmetrical Latency Requirement

Asymmetrical latency can pose problems to various applications. This issue can happen during network congestion, when one way is slower than the opposite way. Another scenario is asymmetric routing. In this case, incoming and outgoing routes take different paths. This can cause issues especially in PTP (Precision Time Protocol), which requires consistent latency. As was tested in [4], asymmetrical latency can temporarily decrease the protocol accuracy from 1 μs to up to 168 μs. Correct time synchronisation is especially important for smart grid devices such as IEDs and PMUs (Phasor Measurement Units).

4 SDN Compliance to the Latency Requirements

Two experimental networks were constructed to complexly verify SDN compliance to analysed smart city latency requirements. The first scenario represented a **small scale physical network**. It was composed from two interconnected HP 3800 switches (KA.15.17.0008), each with two connected hosts (OpenStack virtualized Ubuntu-based machines with 8 GB dedicated RAM). The goal of using hardware devices was to verify particular performance issues, which could not have been reliably tested in emulated environments.

The second scenario represented an **emulated large scale network** corresponding to a topology of a real smart city. The size of the core network (using Open vSwitches) was inspired by Bristol smart city[1] (the number of IoT nodes—virtualized hosts with shared Ubuntu kernel—was reduced). The goal of this network was to verify specific smart city latency issues, which could be influenced by the number of network hops. The topology (displayed in Fig. 1) was setup in the Mininet emulator (2.2.1), which run on the following host machine:

- CPU—AMD Ryzen 7 1800X 8-Core, 3.60 GHz
- RAM—32 GB, 2400 MHz
- Storage—Samsung SSD 850 Pro, 256 GB
- OS—Windows 10 (Host), Ubuntu Mate 16.04.3 (Guest)

[1] More information about the project: https://www.bristolisopen.com/.

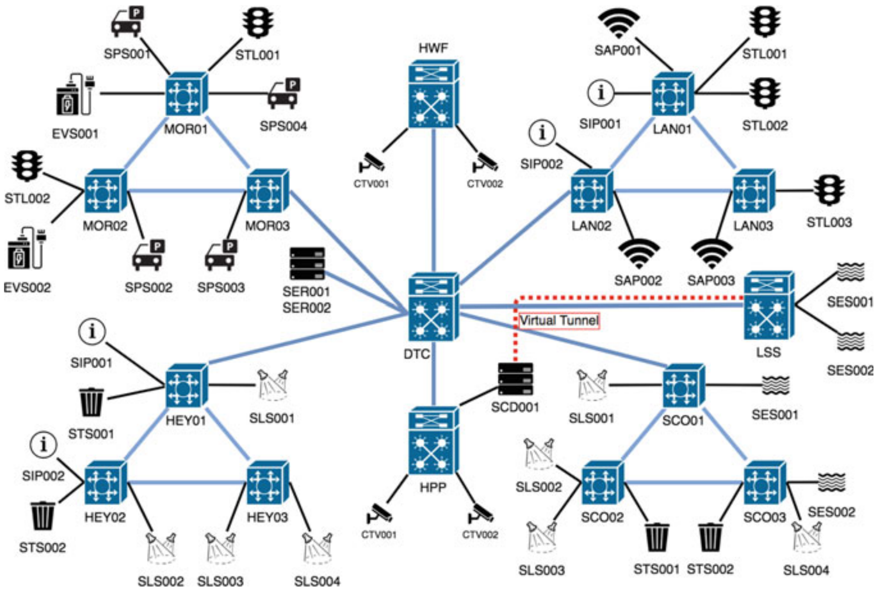


Fig. 1 Emulated network topology

In both cases, the networks were managed by the RYU controller with the default module *simple_switch_13.py*.

4.1 Methodology of Latency Testing

Measuring latency requires use of a traffic capturing tool to compare timestamps of messages on appropriate interfaces. The *Wireshark* tool was chosen for this purpose. The measurement also requires precise time synchronisation between end devices. From this perspective, the best results can be achieved in the Mininet emulator, because the emulated devices share the same OS kernel and therefore have the same internal clock. On the other hand, two-way latency (RTT) can be easily measured with the common *ping* tool even in the real network.

4.2 Typical SDN Latency

A typical SDN latency was measured in two scenarios within the emulated network. In the first scenario, the traffic was sent between *SAP001* and *SER001* (shorter path via 3 hops). In the second scenario, traffic was sent between *SAP001* and *STS002* (longer path via 5 hops). The results from the three first ICMP messages in both experiments are shown in Table 1.

Table 1 Typical SDN latency

Source—Destination, packet n.	Request L. (ms)	Reply L. (ms)	RTT (ms)
SAP001_LAN—SER001_DTC, 1.	0.451	0.166	0.654
SAP001_LAN—SER001_DTC, 2.	0.024	0.006	0.077
SAP001_LAN—SER001_DTC, 3.	0.022	0.006	0.069
SAP001_LAN—STS002_HEY, 1.	0.666	0.268	0.979
SAP001_LAN—STS002_HEY, 2.	0.026	0.009	0.075
SAP001_LAN—STS002_HEY, 3.	0.027	0.010	0.076

In both cases, the OpenFlow rules were already located in flow tables of all affected forwarding devices. Therefore, an increased latency of the first packet from each flow, is caused only by ARP, which translates the IP address to the MAC address. Both first packet latencies of the longer path are consistently higher than the latencies of the shorter path. Additionally, RTT was measured and it corresponded to the sum of request latency, reply latency, and processing time of the queried device.

4.3 Latency of Proactive and Reactive Flow Rule Insertion

This test verified latency of proactive flow rule insertion (used on request) and reactive flow rule insertion (used under normal circumstances). Due to required time synchronisation, it was also performed in the emulated network. In the first scenario, a path was created by pre-populating OpenFlow rules into the affected devices. The path corresponded to the virtual tunnel depicted in Fig. 1. The source device was *SES001* and the destination was *SCD001*. Additionally, the IP-MAC mapping for the specified traffic was manually inserted into both end devices by the following topology script modification:

```
scd001_hpp.cmdPrint('arp -s 10.7.8.1 00:00:00:00:00:61')
ses001_lss.cmdPrint('arp -s 10.8.10.1 00:00:00:00:00:73')
```

In the second scenario, the path was not created, but the manual IP-MAC mapping was still set. The results of both scenarios are shown in Table 2.

The latency and RTT are significantly lower in the case of proactive flow rule insertion. As RTT suggests, the biggest delay in reactive flow rule insertion is caused by controller processing. Additionally, the reactive flow rule insertion latency can significantly increase in a more complicated network topology. The increase will depend on the controller approach to handle reactive flow rule insertion. There are two general methods:

- Sequential—the easiest approach implemented in simpler controllers such as RYU. It responds to all *packet-in* messages received sequentially by all forwarding

Table 2 Latency of proactive and reactive flow rule insertion

Source—Destination, packet n.	Request L. (ms)	Reply L. (ms)	RTT(ms)
Proactive flow rule insertion			
SES001_LSS—SCD001_HPP, 1.	0.200	0.1	0.34
SES001_LSS—SCD001_HPP, 2.	0.014	0.003	0.057
SES001_LSS—SCD001_HPP, 3.	0.014	0.002	0.056
Reactive flow rule insertion			
SES001_LSS—SCD001_HPP, 1.	0.493	0.0139	9.52
SES001_LSS—SCD001_HPP, 2.	0.001	0.0004	0.069
SES001_LSS—SCD001_HPP, 3.	0.002	0.0003	0.068

devices along the path. This therefore linearly increases the first packet latency according to the number of forwarding devices.

- Instant—a more advanced approach implemented, for example, in the Floodlight controller. It sends *flow-modification* messages to all corresponding devices along the entire path at once—in reaction to the first *packet-in* message. The first packet latency should therefore not significantly vary in different network topology sizes.

It is important to mention, that the increased latency applies only to the first message of a flow. Latency of the following messages is similar regardless of the used rule insertion method. The latency could again increase only if a flow rule would expire and a new one would have to be inserted.

4.4 *Effects of HW Flow Table Implementations on Latency*

Different implementations of hardware flow tables used in real forwarding devices, can have various effects on latency. Flow tables on hardware devices can be implemented in several data structures, which most often utilise existing specialised chips such as ASICs (Application Specific Integrated Circuit) and TCAM (Ternary Content Addressable Memory) tables. These tools should provide consistent lookup times, no matter how full the table is (even if the table is almost full, lookup times should be the same as when the table is almost empty). The purpose of this test was to verify this assumption. The real network had to be used to utilise the hardware, so only the RTT was measured.

Firstly, only several flow rules were inserted into the flow table of *switch 1* and RTT between node 1 and 2 was measured. Then a *script* was used to generate *general* OpenFlow rules until the table was completely full. The performance was tested again. In the last scenario, the script was modified to generate more *specific* rules. These rules contained source and destination MAC addresses, VLAN ID, IP protocol, and source and destination TCP ports. The following code was used to create the flow match:

Table 3 HW flow table implementation and RTTs

Experiment description	Avg. RTT (ms)	Max. RTT (ms)
(1) Empty table (few rules)	0.604	1.083
(2) Full table, general rules (2 match fields)	0.688	1.283
(3) Full table, specific rules (7 match fields)	0.526	0.780

```
match = parser.OFPMatch(eth_src=rand_mac_src, eth_dst=rand_mac_dst,
eth_type=0x0800, vlan_vid=10, ip_proto=6,
tcp_src=rand_tcp1, tcp_dst=rand_tcp2)
```

Average results from 10 experiments are summarised in Table 3. The assumption was proven correct—there is no significant difference in measured RTTs.

4.5 Processing Delay of the SDN Controller

Various SDN controllers can have different performance. Moreover, this performance depends on the server on which the controller is running, on server's utilisation, and on the network topology. This scenario tested, how well can the RYU controller cope with strict latency requirements. The goal was to measure, how long it takes the controller to react on a network event. The real network was used for the test and the controller was deployed in a virtualized Ubuntu-based device (with 8 GB of RAM) running in OpenStack. Data traffic was sent between node 1 and 2 and time between *packet-in* events and *flow-modification* messages was measured on the controller's interface. In total, ten experiments were conducted and the average measured latency was: **4.491 ms**.

The effect of this delay is apparent in impact on response time. *Ping* was used to generate ICMP (Internet Control Message Protocol) messages between the same two devices (node 1 and 2) and the response times were captured. As expected, the first packet of every flow had significantly higher RTT than the following ones. Average RTT value from 10 measurements for the first packet was: **17.2 ms**, for the second packet: **0.582 ms**, and for the third packet: **0.668 ms**.

The increased RTT of the first packet is caused by processing on the controller. The controller must decide in software, where to forward the packet and consequently insert a flow rule into the appropriate device. The subsequent messages are therefore handled by this rule immediately on the switch and preferably in a hardware table (although a software table can be used as well) thus significantly reducing the processing time.

4.6 Discussion

Tests of the most common type of the flow rule insertion method (reactive) performed on real devices, measured the average first packet RTT to be 17 ms. This approximately corresponds to **8.5 ms** one-way latency (assuming symmetric latency). Results in the emulated network showed 9.5 ms and therefore **4.8 ms** for symmetric one-way latency. This clearly exceeds requirements of the most demanding smart city applications: IED internal communication (3 ms) and 6LoWPAN connection establishment (5 ms). Furthermore, the results are on the edge with information protection communication (8 ms). These types of communication therefore require proactive insertion method. Measured latency of this method was **0.3 ms** in the real network and **0.2 ms** in the emulated network. This approach of flow rules pre-loading has disadvantages in higher flow table utilisation and a more complicated application logic. An additional lowering of the latency can be achieved by configuration of static IP-MAC mapping. This will eliminate extra latency caused by ARP (IPv4) or Neighbor Discovery protocol (IPv6).

The remaining types of communication can be accomplished with the less demanding reactive insertion method. Even some latency demanding applications like game streaming can be safely operated using the slower reactive method. This is based on the fact, that only the first packet (processed by the controller in software) will have increased latency, while the following packets will be handled in hardware (by forwarding devices). The nature of these applications does not require minimal latency at all times (for example during a game initialisation). On the other hand, some critical applications like node connectivity, or substation control and monitoring, should use the proactive method even while their latency requirements could be met with the reactive method. This recommendation is based on the fact, that a sudden network load could increase the latency behind specified limits. Clearly, the

Table 4 Smart city latency requirements and SDN

Application	Domain	Required latency (ms)	SDN[a]
IED communication	SG	3/10	P
Nodes connectivity (6LoWPAN)	All	5 (not critical)	P
Information protection	SG	8	P
Games streaming	SB	10	R
Nodes connectivity (Zigbee)	All	15 (not critical)	P
Substation monitoring and control	SG	16	P
Safety-related communication	SM	50	R
Cloud applications	SB	50	R
Non safety-related communication	SM	100	R
VoIP and video conferencing	SB	150	R

[a]Recommended proactive/reactive flow insertion method

proactive method in these cases can be safely used only if the capacity of affected flow tables is sufficient.

The summary of these recommendations is displayed in the last column in Table 4.

5 Conclusion

The performed testing verified, that SDN can be safely utilised in current smart city scenarios. Most of the time, the SDN impact on latency, when compared with traditional networks, is negligible. This is achieved by hardware implementation of flow tables, which utilise existing chips of forwarding devices. In specific conditions such as a new connection establishment, expired flow rules, and need of software packet processing, the latency of SDN can be above the smart city requirements. In these cases, proactive flow rule insertion has to be used to achieve the below-threshold latency requirements. It is important to mention, that all the measured data represents a network without congestion. In a real world scenario, such a network has to be designed with sufficient over-provisioning in mind. In traditional networks, an average link utilisation should be maximally 40%. If the network becomes congested, the latency can further increase. Similar increase can occur in a more complicated network topology, or by usage of a slower controller, server, or forwarding devices with lower performance. All these conditions can result in the need to use proactive flow rule insertion even for less demanding applications.

References

1. Celino, I., Kotoulas, S.: Smart Cities [guest editors' introduction]. IEEE Internet Comput. **17**(6), 8–11 (2013). https://doi.org/10.1109/MIC.2013.117
2. Cleveland, F.: IEC TC57 WG15: IEC 62351 security standards for the power system information infrastructure. Security Standard ver 14, International Electrotechnical Commission (2012)
3. Demmel, S., Lambert, A., Gruyer, D., Rakotonirainy, A., Monacelli, E.: Empirical IEEE 802.11p performance evaluation on test tracks. In: 2012 IEEE Intelligent Vehicles Symposium, pp. 837–842 (2012). https://doi.org/10.1109/IVS.2012.6232268
4. Endace: IEEE 1588 PTP clock synchronization over a WAN backbone. Whitepaper, Endace Technology Limited (2016)
5. ETSI SG Group: Smart Grid Reference Architecture. Technical Report. Mandate M/490 v3.0, CEN-CENELEC-ETSI Smart Grid Coordination Group (2012). http://gridscientific.com/images/Smart_Grid_Reference_Artichtecture.pdf. Accessed 31 Aug 2017
6. Filippi, A., Moerman, K., Daalderop, G., Alexander, P.D., Schober, F., Pfliegl, W.: Ready to roll: Why 802.11p beats LTE and 5G for V2x. White paper, NXP Semiconductors, Cohda Wireless, and Siemens (2016). http://www.eenewsautomotive.com/design-center/why-80211p-beats-lte-and-5g-v2x. Accessed 28 Nov 2017
7. Ge, B., Zhu, W.T.: Preserving User Privacy in the Smart Grid by Hiding Appliance Load Characteristics, pp. 67–80. Springer International Publishing, Cham (2013). https://doi.org/10.1007/978-3-319-03584-0_6

8. IEEE: IEEE Std 1646-2004—IEEE standard communication delivery time performance requirements for electric power substation automation. Ieee standard, IEEE (2005). https://standards.ieee.org/findstds/standard/1646-2004.html. Accessed: 17 Jan 2018
9. Kailas, A., Cecchi, V., Mukherjee, A.: A survey of communications and networking technologies for energy management in buildings and home automation. J. Comput. Netw. Commun. (2012). https://doi.org/10.1155/2012/932181
10. Li, B., Mirhashemi, M.S., Laurent, X., Gao, J.: Wireless access for vehicular environments. Final report, Chalmers University of Technology (2011). https://pdfs.semanticscholar.org/8702/4b92ac23dc6e1a4322f760b77e93914826ae.pdf. Accessed 8 Jan 2018
11. Mulla, A., Baviskar, J., Khare, S., Kazi, F.: The wireless technologies for smart grid communication: a review. In: 2015 Fifth International Conference on Communication Systems and Network Technologies, pp. 442–447 (2015). https://doi.org/10.1109/CSNT.2015.146
12. Phan, A., Qureshi, S.T.: 5G impact on Smart Cities (2017). https://www.researchgate.net/publication/315804922_5G_impact_On_Smart_Cities. Accessed 18 Feb 2018
13. Wang, W., Xu, Y., Khanna, M.: A survey on the communication architectures in smart grid. Computer Networks **55**(15), 3604–3629 (2011). https://doi.org/10.1016/j.comnet.2011.07.010
14. Zigbee Alliance: Zigbee resource guide. Resource guide, Zigbee Alliance (2017). http://www.nxtbook.com/nxtbooks/webcom/zigbee_rg2017/. Accessed 5 Jan 2018

Speech Activity Detection for Deaf People: Evaluation on the Developed Smart Solution Prototype

Ales Berger and Filip Maly

Abstract This research constitutes a relatively new approach by developing a smart solution which has emerged from the research activity using at first Google Glass and usable speech detection services. The authors conducted in the last year a series of developing, testing and evaluating the prototype results in order to decide, which service provides better results than the third-party speech detection service like Google Speech API or IBM Watson Speech To Text. This finding should significantly help the authors during the data evaluation and testing in developed smart solution. The basic idea is that authors have already developed a functional basic solution—a prototype. This solution was properly working and usable, but there are still some disadvantages to be improved. In order to accomplish the best results possible, the authors have added another element to their solution. A challenging problem which arises in this domain is concerned with significant data savings, server load, detection quality, and again opens a space for further improvements, such as following research and testing. This element is part of the statistical analysis and it is called Hidden Markov Model, which is used for speech recognition applications for last twenty years. The authors examined and studied many different articles and scientific sources in order to find the best solution for higher efficiency of speech recognition usable in their developed prototype (and for this article).

Keywords Speech and natural language processing · Voice detection · Smart device · Smart solution · Android OS · Deafness · R language · RStudio

A. Berger (✉) · F. Maly
Faculty of Informatics and Management, University of Hradec Kralove, Hradec Kralove, Czech Republic
e-mail: ales.berger@uhk.cz

F. Maly
e-mail: filip.maly@uhk.cz

M. Huk et al. (eds.), *Intelligent Information and Database Systems: Recent Developments*, Studies in Computational Intelligence 830,
https://doi.org/10.1007/978-3-030-14132-5_5

1 Introduction

The first chapter of the paper focuses on the basic definitions of the key words and terms necessary for this research. Today, speech-controlled applications and devices that support human speech communication become more and more popular. With the use of mobile devices, availability is no longer limited to a certain place; instead, it is possible to communicate in almost any situation. Efficient and convenient human-computer interfaces based on speech recognition allow us to control devices using spoken commands and to dictate text [1].

Speech and natural language detection and its processing are currently used in many different fields of science, but also in interesting industrial applications. Speech activity detectors take place during speech recognition or speech enhancement algorithms. Other applications may be in the field of communication. Quite similar approach provides VoIP, where speech activity detectors serve to reduce the required transmission capacity.

A series of recent studies has indicated that speech activity detectors can contribute to increase the speech recognition efficiency. The higher efficiency is accomplished by not transferring silent packets. Spoken speech recognition is divided into two steps working in the developed prototype. The first step takes place on the client and the second part on the server. Prior to implementing the server section, the authors performed an analysis of publicly available audio processing technology. There are many ways to detect speech activity and spoken language. The process of detection of specific words handles the Google Speech API, which addresses the problem correctly and efficiently. For authors, VAD (Voice Activity Detection) is interesting to detect speech activity. Authors have tested this principle in the article by using the Hidden Markov Models method implemented in a smart device.

The client part will use the principle of advanced statistical method called Hidden Markov Models (abbreviated as HMM). This paper describes how to proceed and how to use the potential of HMM during development issues. The whole implementation of testing takes place in the R language and in the RStudio environment.

The authors examined and studied many different articles and scientific sources in order to find the best solution for higher efficiency of speech detection and recognition usable in their developed prototype. However, each method, model or technology is diverse and working differently than the authors want or need.

2 Voice Activity Detection

Voice activity detection (VAD) is a method to discriminate speech segments from input noisy speech. It is an integral part to many speech and audio processing applications and is widely used within the field of speech communication for achieving high coding efficiency and low bit rate transmission. Examples include noise reduction for digital hearing aid devices, mobile communication services, voice recognition systems, compression, and speech coding [2]. Voice activity detection usually addresses a binary decision on the presence of speech for each frame of the noisy signal. Approaches that locate speech portions in time and frequency domain, such as speech presence probability (SPP) or ideal binary mask (IBM) estimation, can be considered as extensions of VAD that exceed the scope of this article.

Most of the algorithms proposed for VAD are divided into two processing stages:

- First, features are extracted from the noisy speech signal to achieve a representation that discriminates between speech and noise.
- In a second stage, a detection scheme applies to the features resulting in the final decision [1].

3 Hidden Markov Model and Acoustic Model

Third chapter introduces the definition of advanced statistical method—Hidden Markov Model. Second half of this part is dedicated to Acoustic Speech Recognition. The Hidden Markov model is a finite set of states, each of which is associated with a (generally multidimensional) probability distribution. Transitions among the states are governed by a set of probabilities called transition probabilities. In a particular state an outcome or observation can be generated, according to the associated probability distribution. It is only the outcome, not the state visible to an external observer and therefore states are "hidden" to the outside; hence the name Hidden Markov Model [3]. Hidden Markov models are a general statistical technique for linear problems like sequences or time series and been widely used for speech recognition applications for twenty years.

Because the author's research is dealing with VAD, there is another model suitable for utilization and introduction. An acoustic model is a file that contains statistical representations of each of the distinct sounds that makes up a word. Each of these statistical representations is assigned a label called a phoneme. The English language has about 40 distinct sounds that are useful for speech recognition, and thus we have 40 different phonemes. An acoustic model is created by taking a large database of speech (called a speech corpus) and using special training algorithms to create statistical representations for each phoneme in a language. These statistical representations are called Hidden Markov Models ("HMM"s). Each phoneme has its own HMM [4].

4 Human Ear Limitations and Problems of Speech Recognition

Over 5% of the world's population—360 million people—has disabling hearing loss (328 million adults and 32 million children). Disabling hearing loss refers to hearing loss greater than 40 decibels (dB) in the better hearing ear in adults and a hearing loss greater than 30 dB in the better hearing ear in children. Most people with disabling hearing loss live in low- and middle-income countries. As for elderly people, approximately one third of people over 65 years of age are affected by disabling hearing loss [5]. Due to the ageing phenomenon in a global society, the total number of elderly people affected by hearing loss will be increasing in the future.

A healthy human ear is able to perceive sounds in the range of frequencies from approximately 20 Hz to 20 kHz and in the intensity range of about 0 to 130 dB (the so-called hearing area). However, normal speech lies in a much narrower band—approximately between 180 Hz to 6 kHz and 30 to 80 dB. Good clarity means (in the speech transmission) in the band from 300 Hz to 5 kHz. For the more common conversation, even bandwidth from 300 to 2.5 kHz is sufficient [6].

Speech recognition, or more commonly known as automatic speech recognition (ASR), is the process of interpreting human speech in a computer. A more technical definition is given by Jurafsky [7], where he defines ASR as the building of system for mapping acoustic signals to a string of words.

The key problem with the VAD is associated with a process of speech recognition. A task of speech recognition is confronted with many problems that arise from the nature of human speech. In the moment when a person or human speaks, never pronounce the word twice (or more times) in the same way. This issue is influenced by many various factors. These factors are: emphasis on words, different accents, different pronunciations at the beginning and at the end of the sentence, different melodies of different types of sentences (notification/interrogation), the influence of the previous and the following words in the sentence, emotions, mood, or health condition of the person. Problems can also occur during the process of determining the boundaries between individual words. One does not pronounce words as distinct, separate units; on the contrary, the words closely follow and, for example, non-salient prepositions merge with another word altogether [7]. The physical and psychological impacts of deafness are quite different for every individual impaired person. Deaf people who are without hearing from birth has other problems, then the elderly who struggle with hearing loss in the last phase of their life. These elderlies have spent the whole life with normal hearing, so the loss of hearing can be much more stressful and because of it, they must change their daily living. Because of hearing impairment, impaired person is not able to fully understand the sounds of speech and to hear other sounds from his surroundings [8].

5 Developed Smart Solution

The fifth chapter provides useful insight on the testing phase, which was crucial for this research in order to improve the developed prototype. Firstly, readers need to know more about the current state of the prototype using selected Google Cloud Speech API. Secondly, the authors provide more information about the whole process of testing, evaluating and its outputs.

5.1 Current State of Prototype

A year ago, the authors started their research, firstly using a Google Glass and then continued their research work without GG, by developing a smart solution based on their findings, which is properly working and ready to use. However, the testing and implementing phase showed that the developed solution and its efficiency could be better and improved. The authors' goal is to improve their prototype in order to be simpler and less inefficient. The developed prototype is dealing sufficiently with speech recognition thank to the Google Cloud Speech API. The authors chose Google Cloud Speech API because of its simplicity. Additionally, during the testing phase, this API showed great results and proved to be well-established for the developed prototype. In order to compare the prototype functions with other usable services, the authors focus on the IBM Watson Speech To Text. The authors want to use the results of the comparison in their own solution (i.e., their own service that can replace Google Speech API or IBM Watson Speech To Text). The improvement of author's own service is based on the better speech recognition. This means that the prototype through the "new" service gets only sections of audio containing 100% spoken language (speech or voice). This added functionality provides significant savings of time, server and financial costs.

5.2 First Testing of Prototype and Results

At the beginning of 2018, the authors designed the prototype architecture or smart solution. This architecture is improved general platform usable for any issue (i.e., assistive technology for blind people [9]). This process was fulfilled during the first pre-testing stage in a functional solution (i.e., prototype architecture is communicating with a server and the information are presented to a user in a satisfactory amount of time. After this goal, authors continued to the key (2nd) phase of the research, which is another real-use testing by deaf or hearing-impaired people. Finishing 2nd phase of research showed quite interesting results. These results are based on the experience, tracking and interviewing about 10 deaf (or hearing impaired) people from the Czech Republic. Most of these respondents were satisfied with the results, functioning and the whole idea leading towards their higher independence in communication with a normal hearing people without needing of a transcriber or an assistant. Data were obtained from the interviews and practical experiment that were carried out as part of a larger research project. The first phase of testing proved that the prototype architecture designed by the authors is functional; however, there are still some aspects to be improved. The most significant pro is that authors' solution can recognize the voice (even with a different pronunciation or accent in a common, noisy environment). Second positive output of the solution is associated with the help of neural network, which is converting speech into a text. Testing was conducted on

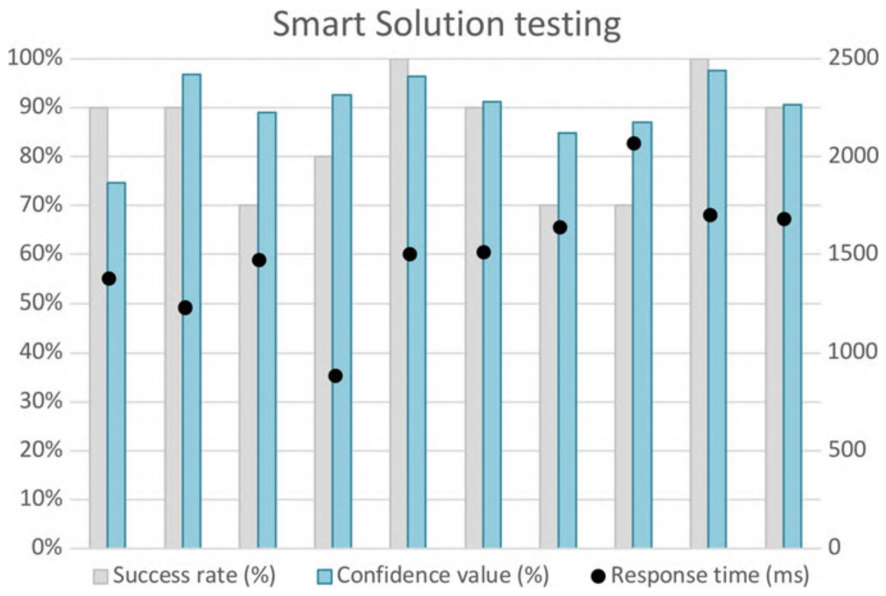


Fig. 1 Results conducted by author's last testing [10]

a smart phone interconnected with a prototype through Wi-Fi. For detailed results, please see Fig. 1.

The most significant result is that in 85% (85 test cases) the designed solution was successful. Only 15% of cases are wrong, it means that the solution is incorrect in 15 test cases from the total number of 100. This success rate of developed solution is quite high and provides very promising results for further research activities. Indicator called "Confidence value" reached a mark 90% for this phase of testing. The average response time is 1515 ms (1st phase of testing), which is not ideal. That is the reason why, the authors are working on improvements (how to save some time) in order to get close to real-time response. The only disadvantage, the users were mentioning, is a little bit longer time to get the output. Due that the important task to work on in further research on this feature of the developed solution [10].

6 HMM Testing

After the first part of testing, the authors decided to explore more usable methods available. The problem is defined as to: "easily find out, if the particular audio is human speech or not". The basic assumption of the authors is that the individual words are not considered. Another presumption is about the spoken speech. Spoken speech has a very specific frequency range. So, for a person, it is quite easy just to find out if someone is talking or not. There are many methods, which can help a lot. After a long literature search and debates with experienced scientists/mathematicians, the authors finally decided to use the help of Hidden Markov Models. HMM is part of statistical analysis and is used for speech recognition applications for more than twenty years. The improved set of testing is based on this statistical method, that is why the authors use RStudio, which is the most suitable development environment.

6.1 RStudio and Its Libraries

The RStudio is an integrated development environment (IDE) for R language. It includes a console, syntax-highlighting editor that supports direct code execution, as well as tools for plotting, history, debugging and workspace management. The RStudio is available in open source and commercial editions and runs on the desktop (Windows, Mac, and Linux) or in a browser connected to RStudio Server or RStudio Server Pro (Debian/Ubuntu, RedHat/CentOS, and SUSE Linux) [11]. There are many functional libraries that is why there is no need and time to develop something new. For testing and verifying that everything works properly, the authors selected some promising libraries already available for RStudio. **The most important libraries for author's R language testing are**:

1. **audio: Audio Interface for R**—interfaces to audio devices (mainly sample-based) from R to allow recording and playback of audio. Built-in devices include Windows MM, Mac OS X AudioUnits and PortAudio (the last one is very experimental) [12].
2. **tuneR: A Collection of Examples**—library tuneR consists of several functions to work with and to analyze Wave files. In the following examples, some of the functions to generate some data (such as sine), to read and write Wave files (readWave, writeWave), to represent or construct (multi-channel) Wave files (Wave, WaveMC), to transform Wave objects (bind, channel, downsample, extractWave, mono, stereo), and to play Wave objects are used. This library can analyse music and speech, can work with different types of audio files such as mp3, midi and wav) [13].
3. **seewave: Sound Analysis and Synthesis**—provides functions for analyzing, manipulating, displaying, editing and synthesizing time waves (particularly sound). This package processes time analysis (oscillograms and envelopes), spectral content, resonance quality factor, entropy, cross correlation and autocorrelation, zero-crossing, dominant frequency, analytic signal, frequency coherence, 2D and 3D spectrograms and many other analyses [14].
4. **HMM: Hidden Markov Models**—easy to use library to setup, apply and make inference with discrete time and discrete space Hidden Markov Models [15].

Second part of the sixth chapter is focused on the data from the TIMIT Speech Database.

6.2 TIMIT Speech Database

The authors decided to use audio recordings (i.e., data) from the TIMIT database. The TIMIT database contains many speeches from many speakers, with approximately 5400. From this database, 300 different speeches were selected from as many speakers as possible, equally represented by both women and men [16]. The main advantage of the TIMIT database is the fact, that the database itself contains many different speeches from many diverse speakers.

TIMIT Acoustic-Phonetic Continuous Speech Corpus (LDC) is a speech-readable database that is primarily designed for studying acoustic-phonetic phenomena and for testing automatic speech recognition systems. With its volume, TIMIT can be categorized as a corpus with a large number of speakers (usually speeches from more than 50 speakers) [17]. The structure of the TIMIT database corpus is quite specific. The TIMIT corpus transcriptions have been hand verified. Test and training subsets, balanced for phonetic and dialectal coverage, are specified. Tabular computer-searchable information is included as well as written documentation. Each recording of the TIMIT database is characterized by the following files:

- An audio file in `PCM` format.
- A file with `PHN` extension containing time-phonetic transcription, written in the Arpabet alphabet.
- A text file containing transcription in the `TXT` format.
- A file `LAB` containing word transcription and its definition in the recording.
- A file `WRD` containing a single word definition in the recording.

6.3 Data Preparation

In order to train the HMM model, it is necessary to pass through all the audio files and process them gradually. At the beginning, the authors have defined a short timeframe for which the authors examined whether it contains a speech activity or not. The time frame was set to 20 ms for presented testing conducted by the authors. In the first step of the testing phase, it is necessary to divide each audio track just after 20 ms. With the included `WRD` files, the authors are able to see which part of the word is in that section. After processing all the data available, the authors presented the results in following Table 1.

7 Model and Results

After the first step of data preparation, the authors know which part of the word is in the given time frame (see Table 1—V1 column). If the value (in V1 column) is N/A, it means that in the given section was not found any word. About other columns—V2

Table 1 Results conducted by author’s data preparation

	V1 <chr>	V2 <int>	V3 <int>	V4 <int>
1	NA	16000	16	1
2	NA	16000	16	1
3	NA	16000	16	1
4	NA	16000	16	1
5	NA	16000	16	1
6	NA	16000	16	1
7	NA	16000	16	1
8	she	16000	16	1
9	she	16000	16	1
10	she	16000	16	1

1–10 of 49,160 rows Previous 1 2 3 4 5 6 ... 100 Next

represents bitrate, V3 number of Bits and V4 is the number of audio track. In case that there are more words in one time frame, the word that prevails, in that section is considered. During this testing is not the aim of research to find out what specific word it was, but to find out whether the timeframe is empty or not. Accordingly, the authors define two groups of sections. The first group is **Speech (1)** and the second one is **Silence (2)** (see Fig. 2). In the silence group are included the timeframes in which there is no part of any word. On the other hand, in the speech group are cumulated all the other timeframes.

In the next step testing, the authors calculated the transition matrix [18]. To put it simply, this means finding out the probability of where the next timeframe can move. This case shows that silence category time frame has 7.87% probability to change to speech category time frame in next iteration. For the transition matrix calculated by the authors exclusively for the purposes of the testing phase, please see Fig. 3.

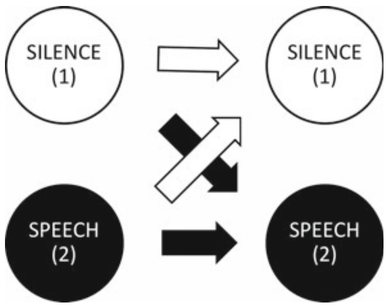


Fig. 2 Two groups defined for author’s testing

```
           [,1]       [,2]
[1,] 0.92131047 0.07868953
[2,] 0.01210323 0.98789677
```

Fig. 3 The transition matrix calculated by the authors

```
$States
[1] "SILENCE" "SPEECH"

$Symbols
 [1]  1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16

$startProbs
SILENCE  SPEECH
    0.5     0.5

$transProbs
        to
from         SILENCE     SPEECH
  SILENCE 0.92131047 0.07868953
  SPEECH  0.01210323 0.98789677
```

Fig. 4 The model preparation calculated by the authors

The next step of the testing phase and research is to categorize parts of the sections into clusters. As a suitable method was selected a *kmeans* method by the authors. A maximum number of iterations were set to 5000. An appropriate number of clusters is 16. These parameters are necessary to define HMM model (please see Fig. 4).

The final step is to implement an algorithm that looks for the most probable sequence of model's states. The authors selected the Viterbi algorithm [19]. This algorithm serves to locate the most probable sequence of states of the model M that generated the observation of P, as well as to calculate its probability. To define the probability of a path by a model that originates from the initial state and ends in state j at time t and at the same time generates the speech vectors $p_1, \ldots, p_t$. Finding the desired sequence of states can be imagined as a search for a path in the grid. On the vertical axis, the individual states of the Markov model are plotted, and the horizontal axes are time slots $1, \ldots, T$ [20].

The resulting chart (see Fig. 5) shows a 9000 ms long audio track with a blue color function, where the most likely located spoken speech in this time frame is (with the highest probability). Second orange function represents time frames with real speech activity. The first set of analyses examined the impact of testing the selected audio track. Due to the utilization of the HMM, the authors proved to save the data transmission more than 30% of tested audio track's length. The presented solution in the RStudio did not skip any time frame, where actually was a spoken speech. The authors assume that the results will not be as successful, when the solution will be deployed in a real-life environment, due significant interferences, noises and other factors. The authors concluded that their prototype solution is very beneficial. Future research should be devoted to the testing and further development of this prototype in order to be used and help people in a real environment.

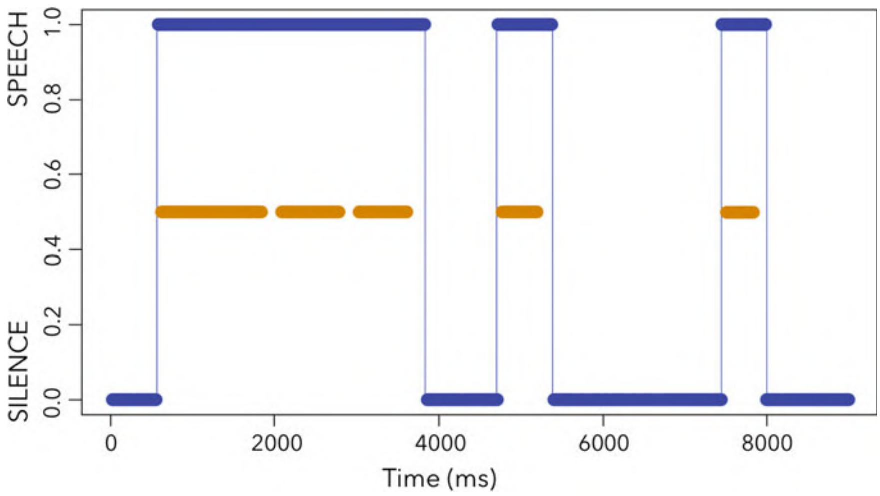


Fig. 5 Resulting chart (of example audio track) calculated by the authors

8 Conclusion

The primary aim of the article is to prove that the selected method is (in its simplicity) usable and brings added value for this research. A secondary goal focuses on the testing phase, which should prove that the selected model could detect somebody speaking or not. That is why the authors selected a well-known statistical method called Hidden Markov Model, which exists quite long and proved great results in speech recognition. No previous study has investigated whether the HMM is used for speech recognition by using a developed solution for a smartphone. The novelty in the developed prototype represents the utilization of a neural network, which poses all the technology needed, saving data transfer and increasing speed of correct results.

In this article, the authors wanted to test the usability (in R language) in order to find out:

1. “Yes, we are going to use it in the developed solution.” Or
2. “It’s not superfluous.”

After testing phase, the authors found out that the HMM “It is really usable.” and their decision is to integrate it in their smart prototype. For the authors and their research, the usability of HMM provides: significant savings of transferred data, savings of a server performance, collection of a new data for further research, opening more possibilities for research of battery savings and another point of view in terms of better speech detection modelling. In the following research, the authors will deploy their solution on the smartphone. After this smartphone deployment the authors can prove that their developer solution (with HMM) is better than the previous one (without HMM).

Acknowledgements This work and the contribution were supported by the project of Students Grant Agency—FIM, University of Hradec Kralove, Czech Republic. Ales Berger is a student member of the research team.

References

1. Graf, S. et al.: Features for voice activity detection: a comparative analysis. EURASIP J. Adv. Sign. Process. **1**, 91 (2015)
2. Yanna, M.A., Nishihara. A.: Efficient voice activity detection algorithm using long-term spectral flatness measure. EURASIP J. Audio Speech Music Process. **1**, 87 (2013)
3. Warakagoda, N.D.: A hybrid ANN-HMM ASR system with NN based adaptive preprocessing. May. Web (1996)
4. Wang, Z., Schultz, T., Waibel, A.: Comparison of acoustic model adaptation techniques on non-native speech. In: 2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003, Proceedings (ICASSP'03), pp. I–I. IEEE (2003)
5. Shearer, A.E., Hildebrand, M.S., Smith, R.J.H.: Hereditary hearing loss and deafness overview (2017)
6. Sohn, J., Kim, N.S., Sung, W.: A statistical model-based voice activity detection. IEEE Sig. Process Lett. **6**(1), 1–3 (1999)
7. Jurafsky, D., Martin, J.H.: Speech and language processing: an introduction to natural language processing, computational linguistics, and speech recognition. Prentice Hall series in artificial intelligence 1–1024 (2009)
8. Kyle, J.G., et al.: Sign language: the study of deaf people and their language. Cambridge University Press, Cambridge (1988)
9. Berger, A., et al.: Google glass used as assistive technology its utilization for blind and visually impaired people. In: International Conference on Mobile Web and Information Systems, pp. 70–82. Springer, Cham (2017)
10. Berger, A., Maly, F.: Prototype of a smart google glass solution for deaf (and hearing impaired) people. In: International Conference on Mobile Web and Intelligent Information Systems, pp. 38–47. Springer, Cham (2018)
11. Gandrud, C.: Reproducible research with R and R studio. Chapman and Hall/CRC (2016)
12. Urbanek, S.: Audio Interface for R. URL: https://cran.r-project.org/package=audio
13. Ligges, U., et al.: Analysis of Music and Speech. URL: https://cran.r-project.org/package=tuneR
14. Sueur, J., et al. Sound Analysis and Synthesis. URL: https://cran.r-project.org/package=seewave
15. Himmelmann, L.: HMM—Hidden Markov Models. URL: https://cran.r-project.org/package=HMM
16. Zue, V., Seneff, S., Glass, J.: Speech database development at MIT: TIMIT and beyond. Speech Commun. **9**(4), 351–356 (1990)
17. Garofolo, J.S. et al. TIMIT Acoustic-Phonetic Continuous Speech Corpus LDC93S1. Web Download. Philadelphia: Linguistic Data Consortium (1993)
18. Aalen, O.O., Johansen, S.: An empirical transition matrix for non-homogeneous Markov chains based on censored observations. Scand. J. Stat. **1**, 141–150 (1978)
19. Lou, H.-L.: Implementing the Viterbi algorithm. IEEE Signal Process. Mag. **12**(5), 42–52 (1995)
20. Tatarinov, J., Pollák, P.: Hidden markov models in voice activity detection. In: COST278 and ISCA Tutorial and Research Workshop (ITRW) on Robustness Issues in Conversational Interaction (2004)

Mobile Application Framework for IobT Hydrogen Skin Moisturizing

Bilguun Jargalsaikhan, Ki Yong Kim, Erdenebileg Batbaatar, Kwang Ho Park, Pham Van Huy, Jong Yun Lee and Keun Ho Ryu

Abstract In this paper, we developed mobile application framework which manages the hydrogen skin moisturizing device. We use react native framework for developing the mobile application. It allows us to manage and control the hydrogen skin device. Our hydrogen skin moisturizing device is based on PCB boards. We were connecting to PCB boards by Bluetooth communication methods. React native is a framework that provides the same results as a native application developed by a mobile application. It's just enough to save time and money for converting the one written code into other platforms.

Keywords Mobile application · Bluetooth communication method · PCB board · React native

B. Jargalsaikhan · E. Batbaatar · K. H. Park
Database Laboratory, School of Electrical and Computer Engineering, Chungbuk National University, Cheoungju City 28644, Korea
e-mail: bilguun@dblab.chungbuk.ac.kr

E. Batbaatar
e-mail: eegii@dblab.chungbuk.ac.kr

K. H. Park
e-mail: khblack@dblab.chungbuk.ac.kr

K. Y. Kim
Daywater Co. Ltd., Kyeonggi-do 338–24, Korea
e-mail: yesggy@daum.net

P. Van Huy · K. H. Ryu (✉)
Faculty of Information Technology, Ton Duc Thang University, Ho Chi Minh City 700000, Vietnam
e-mail: khryu@tdtu.edu.vn; khryu@chungbuk.ac.kr

P. Van Huy
e-mail: phamvanhuy@tdtu.edu.vn

J. Y. Lee · K. H. Ryu
Department of Computer Science, College of Electrical and Computer Engineering, Chungbuk National University, Cheoungju City 28644, Korea
e-mail: jongyun@chungbuk.ac.kr

M. Huk et al. (eds.), *Intelligent Information and Database Systems: Recent Developments*, Studies in Computational Intelligence 830,
https://doi.org/10.1007/978-3-030-14132-5_6

1 Introduction

One of the key concepts in electronics is the printed circuit board also we called PCB. The printed circuit board is the most common name but may also be called "printed wiring boards" or "printed wiring cards". Before the advent of the PCB circuits were constructed through a laborious process of point-to-point wiring [1]. The printed circuit or PCB consists of two parts. There are printed plate with copper track and Software. All types of semiconductors, such as resistance, capacitance, microprocessor, and transistor are installed in the printed circuits with copper tracks. These semiconductor elements are interconnected by the copper track. Instead, they are roaming software designed to manage their operations. This will guide the semiconductor elements with voltage and current according to the program [2, 3].

Bluetooth is a telecommunications industry specification operating in an unlicensed frequency band of 2.4 GHz [4]. That is a specification for the use of low power radio communications to wirelessly link phones, computers and other network devices [5] over short distances which is about 10 m or less. In addition, countries use computers, entertainment systems or televisions, music and telephones for connecting each other. Connecting these devices can be wired by cable and radio signal and infrared. How to connect is more and more sophisticated and easier than ever. Bluetooth is a wireless and automatic connection. Our daily life has become very simple and interesting.

Mobile phone software is widely used. Controllable devices are increasing day by day. Recently, cars are being used by stopping, firing and shutting down the car for using mobile devices.

Our mobile application framework is connecting to device for hydrogen skin moisturizing. This mobile application enables users to manage skin moisturizing device based on PCB board. It creates favorable conditions for users to manage devices from their favorite places.

In this research work, we provide a mobile application framework for hydrogen skin moisturizing device that shows a status of sensors information and allows you to control PCB board. The clients can use the mobile application framework to easily view information and control PCB board on Bluetooth technology. This paper we compare native, hybrid and web based mobile application.

2 Related Works

2.1 Web Based Mobile Application

The Web-based mobile application is run with a web browser on your phone, otherwise, that is the Chrome, Firefox, or Safari browser that are used daily by us [6].

Memory does not use any memory on the user's phone or any space [7]. The advantages of Web applications are simple and fast development. However, more

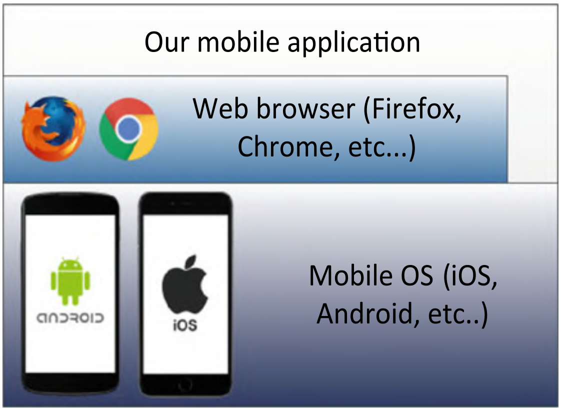


Fig. 1 Architecture of web based mobile application

and more of these weaknesses are: not work in offline mode, not possible to use features of the user's phone, i.e. we do not have access to basic features such as mobile Bluetooth, storage and location [8]. Web based applications are just a great way to deliver presentations and small announcements to many people. The web based mobile application architecture is shown in Fig. 1.

2.2 Native Mobile Application

Native mobile application is a low-level programming and all native applications that we know are written by native apps. In other words, clock, location, contacts, camera, and calendar. However, it means that every platform does not work on clock iOS that is written on Android or used different native apps. The Android application is written in Java programming language, while the iOS application is written in the Objective-C programming language. We can't use the same codebase for each platform and we must write same logic for each one [9].

The advantage of choosing a native app is that it is the fastest and most reliable when it comes to user experience. Native apps can also interact with all of the device's operating system features, such as the microphone, camera, contacts lists, etc. However, a bigger budget is required if you want to build your app for multiple platforms (i.e. iOS and Android) and to keep your native application updated [10, 11]. Additionally, the environment, where we write our code is completely different (Xcode and Android studio). The native mobile application architecture is shown in Fig. 2.

Fig. 2 Architecture of native mobile application

3 Hybrid Mobile Application

Hybrid mobile applications are a very commonly used framework today. Development process is fast and easy for beginners. Web development technologies such as HTML, JavaScript and CSS are used for development of hybrid application.

Developed code is converted to android and iOS. It can be written on one platform and can be used on all platforms, saving a lot of time. However, more resources are used than native applications [12]. Some phone features are not available. The native wrappers allow hybrid apps to be installed on devices, deploy via app stores and access native device APIs via JavaScript [13]. The hybrid mobile application architecture is shown in Fig. 3.

3.1 React Native Mobile Application

In 2015 Facebook launched an open-source project of their own called React Native, which lets you build real, native iOS and Android apps with one codebase. It's not a "mobile web app", a "HTML5 app", or a "hybrid app". With React Native you can build a real mobile app that's indistinguishable from an app built using Objective-C or Java. We just use JavaScript and React to put it together [14]. The react native mobile application architecture is shown in Fig. 4.

In React Native, just as in React, it writes view using JSX, combining markup and the JavaScript that controls it into a single file. For many web developers, the separation of files based on technologies is a given you keep your CSS, HTML, and

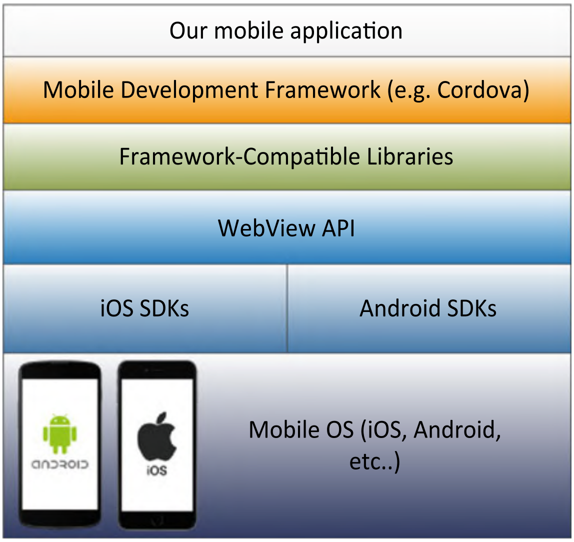


Fig. 3 Architecture of hybrid mobile application

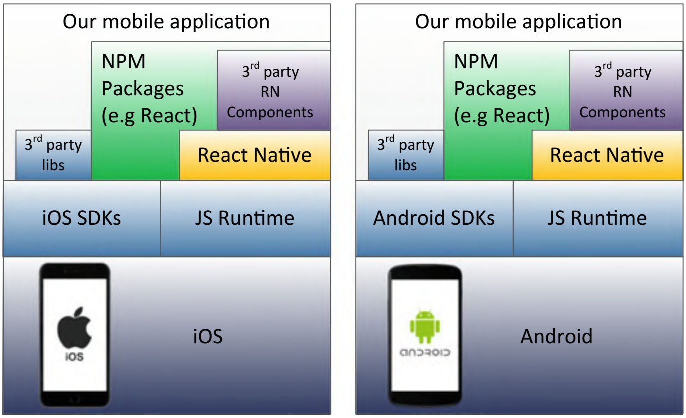


Fig. 4 Architecture of react native mobile application

JavaScript files separate. The idea of combining markup, control logic, and even styling into one language can be confusing.

JSX prioritizes the separation of concerns over the separation of technologies. In React Native, this is even more strictly enforced. In a world without the browser, it makes even more sense to unify our styles, markup, and behavior in a single file for each component. Accordingly, your *.js files in React Native are in fact JSX files. If it was using vanilla JavaScript when working with React for web, it will want to transition to JSX syntax for your work in React Native [15].

4 Proposed Solution

The proposed solution is developing mobile application which is managing to hydrogen skin moisturizing device. The hydrogen skin moisturizing device and connection method architecture is shown in Fig. 5. Our devices connect to other devices using PCB board. Control temp sensor, Sensor Hydrogen, sensor humidity sensors. This control is managed by a user with cell phone software and adjusts the device.

The mobile application is connected to hydrogen skin moisturizing device by Bluetooth communication method. The software will open in two types. The first is to connect new devices. The second is to log your device with access rights. The software will be connected with the hydrogen skin moisturizing device to exchange data and provide users with information such as room temperature, hydrogen, and humidity. Based on the information, the user can change the temp, hydrogen, and humidity settings and turn on and off with the remote control. Our mobile application architecture is shown in Fig. 6.

The mobile application sequence diagram of create device, get status, set hydrogen level and turn on/off of hydrogen skin device is shown in Fig. 7. A sequence diagram shows how the information passed through the user's actions. The user requests to be created the device on the phone software. After that, the software will send a request to connect to the hydrogen skin device and will be displayed to the user if the device is able to connect. If the user wishes to add the device, the device will be retained on the software data.

If the user enters the device status request to the program, the software will send a request to connect the device to the hydrogen skin device. The software will be sent whether the hydrogen skin device is acceptable. If it is available for connect, will send request of status to the hydrogen skin device from the software. When the hydrogen skin device accesses other sensors, it will get the status and send it to the software. The software will show the posted status to users. Based on this, the user will take the following steps.

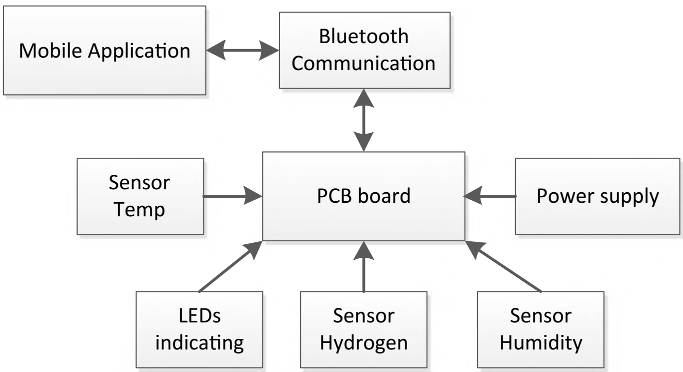


Fig. 5 The hydrogen skin moisturizing device architecture

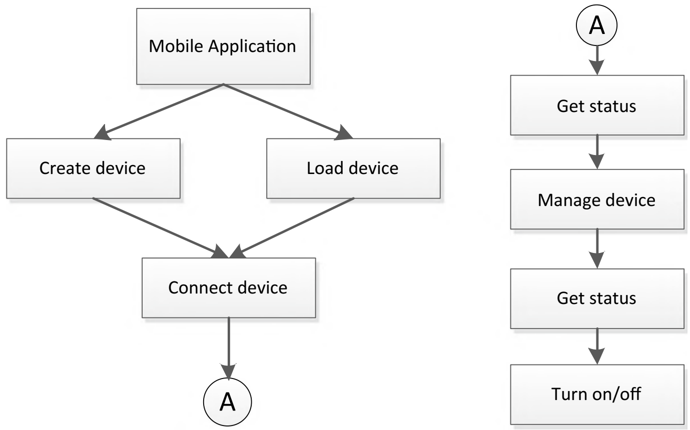


Fig. 6 Mobile application operation

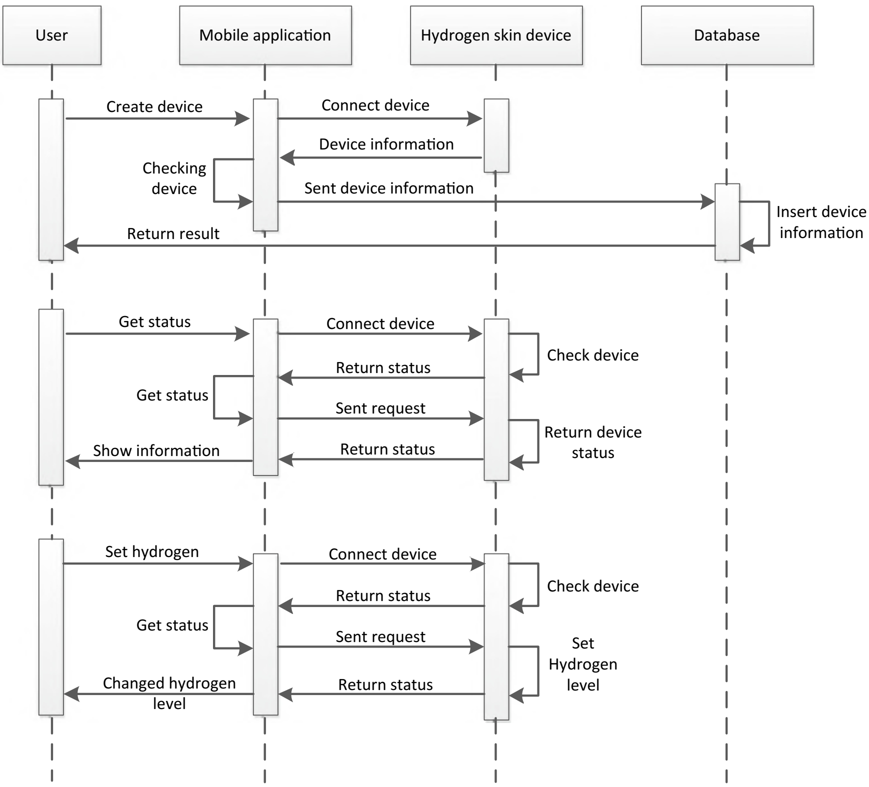


Fig. 7 Sequence diagram of proposed method

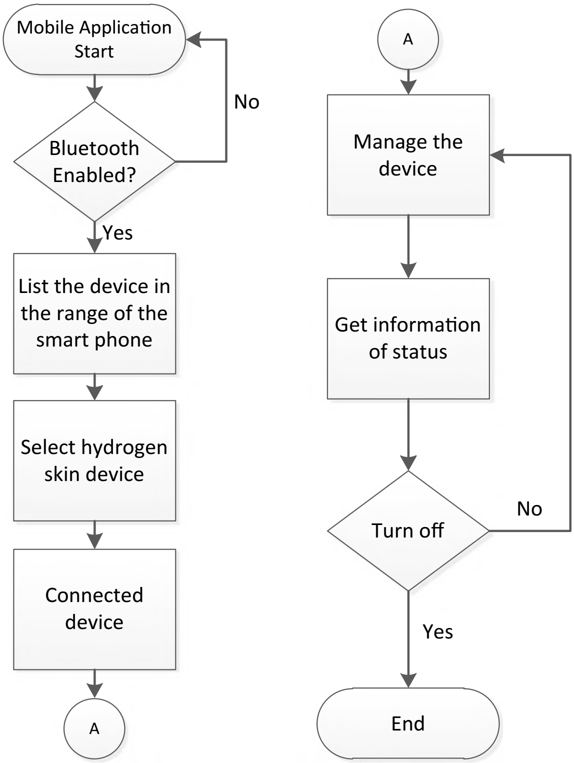


Fig. 8 Flowchart diagram of device connection

The mobile application flowchart diagram is shown in Fig. 8. A flowchart diagram illustrates a mobile application section that connects to the device. Check if Bluetooth is enabled after the software turns on. If Bluetooth is enabled, the list of all hydrogen devices connected and able to connect to the phone will be displayed. After selecting the Hydrogen skin device, it will receive and control the device. If the device is turned off then the operation ends.

5 Implementation

The comparison for mobile application development methods are shown in Table 1. We developed the system on React native mobile application framework.

The mobile application can connect one or more devices and manage all devices. The application can set hydrogen value and time settings. Also, users can view room temperature, humidity and hydrogen. The mobile application's home page is shown in Fig. 9. The device name on the home page will appear on Device 1 and will be moved to the device management page by clicking. If there is no connected device or

Table 1 Comparison table of mobile application

	Web based	Native	Hybrid	React native
Skills needed to reach Android and iOS	HTML, CSS, JavaScript	Objective-C, iOS SDK, Java, Android SDK	HTML, CSS, JavaScript, Development Framework	JavaScript, React
Distribution	Web	App Store/Market	App Store/Market	App Store/Market
Development speed	Fast	Slow	Moderate	Moderate
Development cost	Low	High	Moderate	Moderate
Maintenance cost	Low	High	Moderate	Low
Graphical performance	Moderate	High	Moderate	High
App performance	Moderate	Fast	Moderate	Fast
Access to mobile features	No	Yes	Yes	Yes

change it, click the button at the top right and open the menu. That menu is shown in Fig. 10. This menu includes the keys to enter, register, and select the device and search functions. User login page is shown in Fig. 11. To switch a device to administration page, users must be logged into the system and the user login name and password will be displayed. Also, the user can return from this page. The user registration page is shown in Fig. 12. Fields such as username, password, and email address are included in a form that is shown in the user registered page. In addition, the user can check that the user's name is overlapping with the username in the system before clicking the register button. To do so, click the "Check" button next to the user name field. After completing the registration fields, click the "Next" button to move to the next page. When click on the "Search device" button on the right menu, will move to the page. Click the "search" button located in the center of the screen to get the list of possible devices connected by Bluetooth. Selecting the device name you want to use from this list will be able to manage this device. The Search Device page is shown in Fig. 13. Device management page is shown in Fig. 14. Here are the features for managing devices with the following list. In this regard, the environment humidity indicator (%), moisture adjustment, time settings, settings and "turn off device" button. If the humidity indicator percent that appears in the center of the screen decreases, the color will change and the alert box will appear. The indictor is shown in Fig. 15. By clicking the moisturizing power button, will control the humidifying power of the device at 1–4 levels. The moisturizing power button is shown in Fig. 16. By pressing the time settings button, you can set the time and turn on and off time for the device, and adjust the working time. The setting button is shown in Fig. 17. Click the Settings button to display a window with high, medium and low options. But, pressing the "switch off the device" turns off the device and move to the front of the page.

Fig. 9 Mobile application home page

Fig. 10 Mobile application menu

Fig. 11 Mobile application user login page

Fig. 12 User registration page

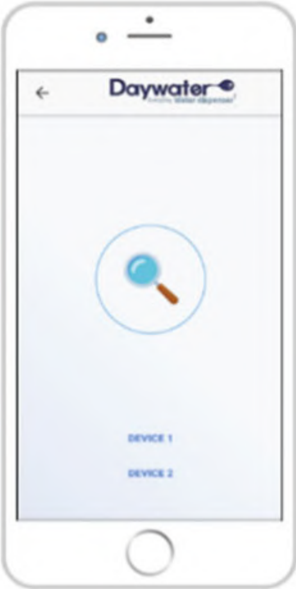


Fig. 13 Search device page

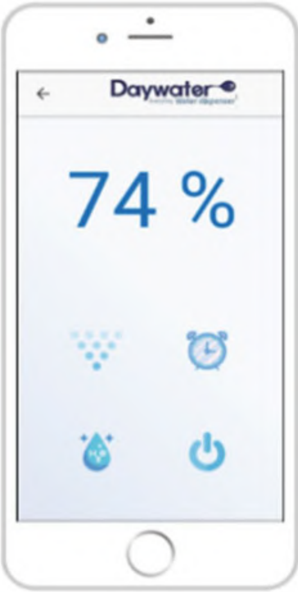


Fig. 14 Device management page

Fig. 15 The humidity indicator

Fig. 16 The moisturizing power

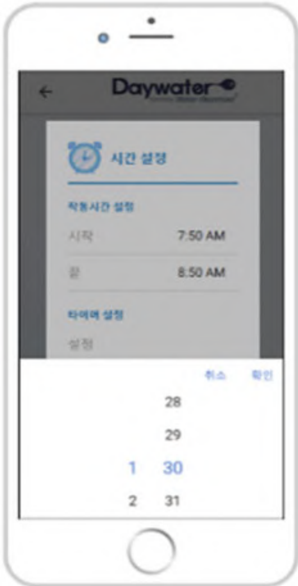


Fig. 17 Time setting page

6 Conclusions

In this paper, we developed mobile application framework which manages the skin device. We use react native framework for developing mobile application.

The mobile application is connecting to hydrogen skin moisturizing device by Bluetooth communication method. The hydrogen skin device is developed on PCB board. The PCB board is installed Bluetooth communication device. Our developed mobile application is working on iOS and Android devices.

In future, we are developing additional functions. For example:

- Get notification on user smart phone
- Get working time (how many time worked?)
- Change the Bluetooth communication method to wireless adapter
- Get alert function on user smart phone

When we used react native framework, there are some advantages in our development. React Native offers a truly native user experience, unlike other hybrid tools that merely provide a native-style wrapper for browser-based applications. It not only makes the development process simple but also faster.

Acknowledgements This work (Grants No. C0541451) was supported by Business for Cooperative R&D between Industry, Academy, and Research Institute funded Korea Small and Medium Business Administration in 2017, by the KIAT (Korea Institute for Advanced of Technology) grant funded by the Korea Government (MOTIE: Ministry of Trade Industry and Energy) (No. N0002429), by supported by the Private Intelligence Information Service Expansion (No. C0511-18-1001) funded by the NIPA (National IT Industry Promotion Agency), by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (No. 2017R1D1A1A02018718).

References

1. Complete PCB design using OrCAD capture and PCB editor, 1st edn, Newnes Publishing, Amsterdam (2009)
2. Cheldavi, A., Nasab, M. H.: Coupling model for the two orthogonal microstrip lines in two layer PCB board (quasi-TEM approach). In: 2005 IEEE International Symposium on Microwave, Antenna, Propagation and EMC Technologies for Wireless Communications, Beijing, Vol. 2, pp. 1096–1099 (2005)
3. https://www.bluetooth.com/bluetooth-technology/topology-options
4. Rashid, R.A., Yusoff, R.: Bluetooth performance analysis in personal area network (PAN). In: International RF and microwave conference proceedings, Putrajaya, Malaysia, 12–14 Sept 2006
5. Bakker, D.M., Gilster, D.M.M., Gilster, R.: Bluetooth End to End, Hungry Minds, Inc., New York (2002)
6. Charland, A., LeRoux, B.: Mobile application development: web versus native. Queue **9**(4), 20 (2011)
7. Hamou, A., Guy, S., Lewden, B., Bilyea, A., Gwadry-Sridhar, F., Bauer, M.: Data collection with iPhone Web apps efficiently collecting patient data using mobile devices. In: 2010 12th IEEE International Conference on e-Health Networking Applications and Services (Healthcom), pp. 235–239 (2010)
8. Terán, M., Aranda, J., Carrillo, H., Mendez, D., Parra, C.: IoT-based system for indoor location using bluetooth low energy. In: 2017 IEEE Colombian Conference on Communications and Computing (COLCOM) (2017)
9. Dalmasso, I., Datta, S.K., Bonnet, C., Nikaein, N.: Survey, Comparison and Evaluation of Cross Platform Mobile Application Development Tools, Mobile Communication Department, EURECOM Sophia Antipolis, France
10. Keur, Christian, Hillegass, Aaron: iOS Programming: The Big Nerd Ranch Guide, 6th edition, Big Nerd Inc, Arizona (2016)
11. Meier, R.: Professional Android 4 Application Development. Wiley, Hoboken (2012)
12. Martinez, M., Lecomte, S.: Towards the Quality Improvement of Cross-Platform Mobile Applications. In: 2017 IEEE/ACM 4th International Conference on Mobile Software Engineering and Systems (MOBILESoft) (2017)
13. Hoehle, H., Venkatesh, V.: Mobile Application Usability: Conceptualization And Instrument Development, Department of Information Systems, Sam M. Walton College of Business, University of Arkansas
14. https://facebook.github.io/react-native/
15. Chinnathambi, K.: Learning React. Addison-Wesley Inc., Boston (2016)

Measurement of Usability for Speech Recognition on Ok Google

Emny Harna Yossy, Wayan Suparta, Agung Trisetyarso, Bahtiar Saleh Abbas and Chul Ho Kang

Abstract As technology develops, Google can accept input not only by typing keywords but also through speech. Speech as input to Google is known as the Ok Google feature. However, many search engine users never pay attention to search engine definitions. So, websites are only limited to projects without further management, and that is what happens to the current website. Therefore, in this study, an assessment of usability is conducted on Google especially in the Ok Google feature using Nielsen Model. The research method used in this research is the quantitative research approach, a method which includes approach and kind of research, population and sample, instrument, data collection technique, and data analyze. Data collection techniques include learnability, efficiency, memorability, errors and satisfaction assessment and Likert scale ranges from 1 to 4. It can be signed into categories of sites with usability values, namely "very difficult to understand," "difficult to understand", "easy to understand" and "very easy to understand". The construct validity is with Orthogonal Iterations with Index reliability of 0.91. The results of this study are the usability testing of Ok Google tends to be significance (α) of 5% with data of 124 is 1.66, feedback statement in the questionnaire is considered valid, then the questionnaire results show $\alpha = 0.91$ which means the measurement of usability on Ok Google Feature is an excellent level. The average usability value obtained by Ok Google in Indonesia is included in the category of usability "easy to understand".

E. H. Yossy (✉) · W. Suparta · A. Trisetyarso · B. S. Abbas · C. H. Kang
Computer Science Department, Binus Graduate Program—Doctor of Computer Science, Bina Nusantara University, 11480 West Jakarta, Indonesia
e-mail: emny.yossy@binus.ac.id

W. Suparta
e-mail: drwaynesparta@gmail.com

A. Trisetyarso
e-mail: atrisetyarso@binus.edu

B. S. Abbas
e-mail: bahtiars@binus.edu

C. H. Kang
e-mail: chkang5136@kw.ac.kr

M. Huk et al. (eds.), *Intelligent Information and Database Systems: Recent Developments*, Studies in Computational Intelligence 830,
https://doi.org/10.1007/978-3-030-14132-5_7

In addition, it was found that Ok Google has the highest memorability level and the lowest error.

Keywords Speech recognition · Website · Usability

1 Introduction

Google website is a search engine site. Users can do data search by typing keywords in Google. Now, Google can accept input not only by typing but also by speech. Technology to search data with speech keywords is known as speech recognition.

Speech recognition takes raw data and acting based on data classification. Speech recognition is a technology that uses equipment with the source of input is speech, like microphones to interpret human voices [1]. Speech recognition is a process that is carried out by a computer to recognize a word spoken by someone regardless of the identity of the person concerned [2]. Speech recognition is also one way for speech to be recognized, so that it can be utilized [3]. One of the uses of speech recognition technology by Google is to provide on Ok Google features. Ok Google uses speech recognition as input voice searches and actions. Ok Google uses speech to do actions like search, get directions, and create reminders. For example, if it is forecast rain, say, "Ok Google, do I need an umbrella tomorrow?" On some devices, you can also say, "Hey Google" [4].

With the addition of speech recognition features as one indicator of the success of the development of the Google website. Google's success is supported by the ownership of a good website interface. The interface of the college website is a bridge that connects user information needs to the sources and services available on Google. Google, which is prepared as a means of searching for customers, needs to be designed, improved and managed properly. To realize broader and more widespread use, open opportunities for accessing, managing and utilizing information on Ok Google in large volumes quickly and accurately, then required a form of evaluation of the website owned by Google, so the quality of websites is known in terms of usability. Assessment of usability is very oriented to how the users perceive when using a product. If users feel the benefits, convenience, efficiency during using time or other positive things, chances are the user will continue to utilize these products and automatically, the usability of the product is at a high level.

Usability is one of the factors that influence the quality of web [5]. A website that has a quality interface that is efficient, effective, satisfy customers, easy to learn, easy to remember and not error [6]. Websites that have a simpler quality interface are effective, efficient and satisfying [7]. However, existing websites need to be evaluated to find out whether Google with the ok feature also has a high-quality interface and obtains user input on the use of college websites. User input is useful for the development of the next website. Site evaluation nowadays, the ease of solution-oriented customer service has become an important business trend. The customer service process must be friendly and easy to use.

However, there is a variety of information, but many of these websites cannot fulfill the initial goals of creation, some even disappoint users. It is known that 60% of the website content is the information that is not needed by the user, and this has an impact on decreasing productivity and increasing frustration [8]. With bad experiences obtained when visiting a website, the user will give a separate assessment or even mark a website as not worth a visit. So many websites have been abandoned by users because of failure to achieve the initial goal of the website. Another disadvantage is that the website profits will decrease and then the website will close.

Forrester Research conducted a study and found that 50% of potential sales were lost because users did not get sufficient info, 40% of those visiting did not return to the website because of the bad experience of the first visit [9]. Although these results do not present all the outstanding websites, at least there is a negative impact when the website is not visited again. To answer this problem, the website that is built must be able to obtain a high level of satisfaction without leaving the specific purpose of the website itself, in this case, the Ok Google. To realize broader utilization and open opportunities for accessing, managing and utilizing tertiary information in large volumes quickly and accurately, a form of evaluation of websites owned by Google is needed so that the quality of the websites seen is known in terms of usability.

Usability comes from the word usable, which can be used properly. Usability has a definition where products are used by certain users to achieve their goals more effectively, efficiently and satisfactorily in the scope of their use [7]. Usability as user experience in interacting with applications or websites until users can operate it easily and quickly [6].

Assessment of usability is very oriented towards how users perceive when using a product. If the user feels the benefits, convenience, and efficiency during the using time or other positive things, the chances are that the user will continue to use the product and automatically the usability level of the product is at a high level. Many users do not like bad website designs and want the website to be easily learned. Therefore, usability measurements from the website are needed. Website measurement consists of four aspects including (1) learnability: the convenience of users in studying the website when they first visit; (2) efficiency: the speed of the user understands the site when he first visits; (3) memorability: ease of users remembering how to use the website when it has not been visited for long; (4) error: ease of handling users when an error occurs when visiting the website; (5) satisfaction.

This study was conducted to assess the usability of speech recognition on Ok Google with control over the usability of the website. Ok Google was created to support Google to use "Ok Google" voice searches and actions. User can use voice to do actions like search, get directions, and create reminders [4].

Researchers also want to know the learnability, efficiency, memorability, errors and satisfaction level of usability in Ok Google and the assessment of efficiency criteria. By using usability analysis from Nielsen, it is hoped that users of this search engine website feature can provide an evaluation of search engine website to increase the use of the website. User satisfaction includes feelings and opinions when visiting the website, which can be described as follows (Fig. 1):

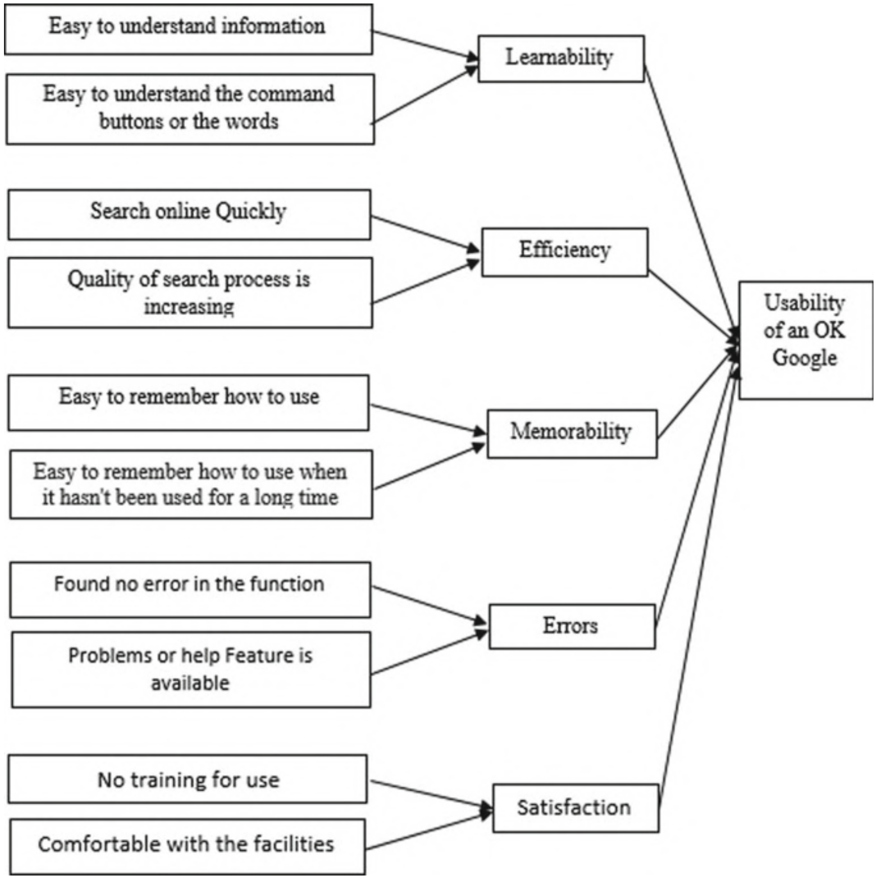


Fig. 1 Nielsen usability model [6]

2 Research Methods

This study aims to explain whether there is a relationship between the speech recognition display on the Ok Google feature and the results of the usability assessment. In this study, there are five research variables namely learnability, efficiency, memorability, error and satisfaction. These variable indicators are developed into the statements stated in the questionnaire using a Likert scale for variable learnability, efficiency, memorability, error, and satisfaction. The next step, data obtained is analyzed using a statistical program. Therefore, this study uses quantitative research methods. Quantitative research is a study to test objective theory by testing the relationships between variables. Variables can be measured using instruments so that the amount of data can be analyzed using statistical procedures.

The research measures the level of speech recognition display in the Ok Google feature with the results of the usability assessment. The resulting data is tested using the Product Moment formula to determine the magnitude of the correlation coefficient and determine whether the relationship between the two variables is significant or not. Research includes the type of correlational research. Correlational research is a nonexperimental type of research where researchers measure two variables and correlations between them with little or no effort to control variables [10].

The population of this research is the search engine users called Google which have the ok Google feature that accepts input using sound. The thing that is considered in determining the population is users who have the potential and enormous contribution to search using Google. The research sample is based on the calculation of the usual testing condition used in statistics (at least 30 subjects). The samples obtained were 36 subjects.

In this study, the data collection tools (instruments) used is non-tests, in the form of questionnaires. The questions or statements in the questionnaire were developed based on usability theory relevant to each research variable. Questions or statements in questionnaires are measured using a Likert scale, namely “a scale used to measure attitudes, opinions, perceptions of a person or group of people about social phenomena”. The answers to each instrument item in this study vary from very positive to very negative, in the form of words such as: strongly agree, agree, disagree, and strongly disagree. Thus, in measuring the research variables, respondents were asked to state their perceptions by choosing one of the alternative answers on a scale of one to four. The development of instruments for each research variable is as follows: (1) learnability: easy to understand website for information and navigation that is presented; (2) efficiency: websites can be accessed quickly and have good quality results; (3) memorability: how to use the website is easy to remember for new user and when the user has not been used it for a long time; (4) error: the website has functions that do not work properly, and the website has problem or help features to overcome errors; (5) satisfaction: users need training or not to access the website and the facilities provided by the website make users feel comfortable.

A questionnaire is to diagnose the level of usability of speech recognition on Google, especially the Ok Google feature. There are ten statements used to measure the usability level of the Ok Google feature. All statements are formulated in the following positive statement sentences: (1) The information provided by Ok Google is easy to understand; (2) User easily understand the command buttons or the words used in Ok Google; (3) User can do online search quickly; (4) User feels the quality of the search process with the application is increasing; (5) It is easy to remember how to use Ok Google; (6) It is easy to remember how to use Ok Google when you have not used it for a long time; (7) User does not find errors or functions that are not working properly in the use of Ok Google; (8) Ok Google provide a Problem or help feature to resolve the problem; (9) User does not feel burdened (do not need to attend training) to access Ok Google; (10) User is comfortable with the facilities available at Ok Google.

Alternative answers given to respond to existing statements include: (1) strongly disagree which means that the activities expressed in the statement are strongly

Table 1 Visually the alternative range of answers and scores

Strongly disagree	Disagree	Agree	Strongly agree
1	2	3	4

Table 2 The following operational planning variables number of statement items, lower total score and highest total score

No.	Usability variable	State number	Total state	Lowest total score	Highest total score
1.	Learnability	1–2	2	2	4
2.	Memorability	3–4	2	2	4
3.	Efficiency	5–6	2	2	4
4.	Error	7–8	2	2	4
5.	Satisfaction	9–10	2	2	4
Total				10	40

disagree to be carried out, the score given is one; (2) disagree which means that the activity expressed in the statement does not agree to be carried out, the score given is two; (3) agree that means the activity disclosed in the agreed statement is carried out, the score given is three; (4) strongly agree which means that the activities expressed in the statement strongly agree to be carried out, the score given is four (Table 1). In this instrument there are ten statements, the lowest total score is 10 (i.e., the result of multiplication between score 1 and the number of statements which is 10); and the highest total score is 40 (the result of multiplication between scores 4 and the number of statements which is 10). Visually the alternative range of answers and their scores can be described as follows:

Because the usability speech recognition variable on the Ok Google feature is divided into five sub-variables, the total score of each sub-variable will vary depending on the number of statements formulated to capture the relevant sub-variable data. In detail of the number of statements, the lowest total score and the highest total score for each type of operational planning are as follows (Table 2):

In this study, the valuation method uses a Likert scale reference benchmark using a scale of 1, 2, 3 and 4. Value 1 will be given if the information listed on the sub-criteria is not available at all on the site. Value 2 will be given if there is information printed on the sub-criteria, but it is difficult to understand. Value 3 will be given if the information printed on the site is easy to understand. Value 4 is given if the information printed on the site is very easy to understand. The purpose of this assessment score is to get a greater difference in the final score between the two sites when compared with the valuation using the values of 0 and 1. In addition, this aims to make the indicators assessed on the framework more detailed. The values of each sub-criterion that exist on learnability criteria are 4, 3, 2, and 1. Assessment aims to make the indicators assessed in the framework more detailed. The assessment of each criterion has similarities, namely the level of understanding, easy to understand, difficult to understand and very easy to understand.

These levels are adjusted to the sub-criteria mentioned earlier, on the criteria of learnability and memorability which are measured in the level of ease, the efficiency criteria are at the level of speed, the criteria for errors are at the level of error, and the criteria of satisfaction are at the level of satisfaction. Assessments for each aspect carried out based on the sum of the assessment results of all sub-criteria which then be divided by the number of sub-criteria available. The maximum value of learnability, efficacy, memorability, errors, and satisfaction is the maximum value obtained in all sub-criteria in each assessment criterion, which is if all indicators get a value of 4. Because the number of sub-criteria of the five criteria is not the same, then the maximum weighting of each criterion is carried out.

$$\text{Learnability} = \frac{\text{Sum value of learnability sub-criteria}}{\text{Maximum value all of learnability sub-criteria}} \tag{1}$$

$$\text{Efficiency} = \frac{\text{Sum value of efficiency sub-criteria}}{\text{Maximum value all of efficiency sub-criteria}} \tag{2}$$

$$\text{Memorability} = \frac{\text{Sum value of memorability sub-criteria}}{\text{Maximum value all of memorability sub-criteria}} \tag{3}$$

$$\text{Errors} = \frac{\text{Sum value of errors sub-criteria}}{\text{Maximum value all of errors sub-criteria}} \tag{4}$$

$$\text{Satisfaction} = \frac{\text{Sum value of satisfcation sub-criteria}}{\text{Maximum value all of satisfcation sub-criteria}} \tag{5}$$

Based on usability values or the results of the five criteria, the sites can be grouped into categories with usability values "very difficult to understand," "difficult to understand", "easy to understand" and "very easy to understand". Then it is defined about the total usability value, then the usability value of the proposed framework is defined as the average of the sum of the five criteria.

$$\text{Usability Value} = \frac{\text{Sum value of criteria (learnability, efficiency, memorability, errors, satisfaction)}}{\text{Maximum sum all value of criteria}} \tag{6}$$

In this study, the weighting is in grades 1–4 according to the results of the assessment on the proposed framework. Explanation of the overall usability value, namely:

a. Usability on a site is categorized as "very difficult to understand" if the average value is in the range of 0–1.0.
b. Usability on a site is categorized as "difficult to understand" if the average value is in the range of 1.01–2.0.
c. Usability on a site is categorized as "easy to understand" if the average value is in the range of values 2.01–3.0.
d. Usability on a site is categorized as "very easy to understand" if the average value of values is in the range of values 3.01–4.0.

3 Results and Analysis

The results of the questionnaire consisted of 10 columns according to the feedback by respondents on the questionnaire, and 124 lines according to the number of respondents. The correlation of attributes in the table is used to describe the characteristics of an individual or object about how the dependent, associate or relationship between attributes [11]. To measure the correlation attribute of the questionnaire results used Pearson Product Moment Correlation Coefficient.

Pearson Product Moment Correlation Coefficient or Linear Correlation Coefficient is a measure of the strength and direction of the relationship between two variables quantitatively linearly [12]. Pearson correlation r assumes that both variables analyzed have a linear relation [13]. Linear Correlation Coefficient is calculated using the following equation [14]:

$$r = \frac{N\sum XY - \left(\sum X\right)\sum Y}{\sqrt{\left(N\sum X^2 - \left(\sum X\right)^2\right)\left(N\sum Y^2 - \left(\sum Y\right)^2\right)}}$$

To speed up the calculation then use Microsoft Excel application because it provides a function to calculate the value of the linear correlation coefficient with the function name CORREL() [15].

The result of the calculation is compared with the coefficient of the Critical Value table of Pearson Correlation Coefficient r, if more count value is greater than the value of the table, then it is considered valid, and vice versa. The value of the Critical Value table of Pearson Correlation Coefficient r for significance (α) of 5% with data of 124 is 1.66. Table 3 is the calculation results.

After each feedback statement in the questionnaire is considered valid, and then the reliability test is performed. For reliability test, this study used Cronbach Alpha. Analog measure uses Cronbach's Alpha to estimate internal consistency [16]. The results of reliability calculations are shown in Table 4.

The classification of the widely accepted approach is as follows Table 5 [17].

Based on the Classification of Reliability Coefficient (Cronbach Alpha) table, the questionnaire $\alpha \geq 0.9$. Therefore, Reliability Coefficient (Cronbach Alpha) on evaluating the use of Speech Recognition for Search on Ok Google Features is at the excellent level of reliability.

Table 3 Calculation of linear correlation

Respondents	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10
T-hit	12.09	7.49	6.36	7.06	8.40	6.70	4.66	4.35	7.77	9.68
T-tab	1.66									
	Valid	Valid	Valid	Valid	Valid	Valid	Valid	Valid	Valid	Valid

Table 4 The results of reliability calculations

Respondents	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10
Total	377	382	384	374	393	376	306	353	388	369
Item variant	0.54	0.57	0.67	0.55	0.56	0.66	1.02	0.60	0.73	0.65
Total item variant	6.55									
Total variant	36.09									
Reliability	0.91									

Table 5 Classification of reliability coefficient (Cronbach Alpha)

Reliability coefficient (Cronbach Alpha)	Comment
$\alpha \geq 0.9$	Excellent
$0.7 \leq \alpha < 0.9$	Good
$0.6 \leq \alpha < 0.7$	Acceptable
$0.5 \leq \alpha < 0.6$	Weak
$\alpha < 0.5$	Unacceptable

After the researcher found that the questionnaire was valid and reliability in accordance with the standards, the next step was to compare criteria and results of measurement to the results of overall usability. The results are then divided into five parts, namely the results of the assessment on learnability criteria, results on efficiency criteria, results on memorability criteria, results on criteria of errors, results on satisfaction and final criteria are the results of measurement of overall usability or total values of the five criteria there is. The following are the results of the assessment based on dimensions:

1. Learnability

The total value for learnability criteria is obtained from the sum of the results in the sub-criteria, the ease of the user to understand the information given and the command button or words. The average value for the sub-criteria ease of understanding of the information given is 3.040323 from the maximum value of each sub-criteria is 4, the average for sub-criteria the user understands the command or words is 3.080645 from the maximum value of each sub-criteria, and the average for the ease sub-criteria. So, the total average value for the learnability criteria of the maximum value of the two sub-criteria is equal to 3.060484. Based on these data, overall it can be concluded that the usability level on Ok Google for learnability criteria is at a good level because the average reaches number 3.

2. Efficiency

The total value for the efficiency criteria is obtained from the sum of the results in the sub-criteria makes it easy for users to search quickly and the quality of the search process increases. The average value for the sub-criteria for the ease of users to do a quick search is 3.096774 from the maximum value for each sub-criteria is 4, the average for the sub-criteria quality improvement process is 3.016129 from the maximum value of each sub-criteria, and the average for the ease sub-criteria. So, the total average value for the efficiency criteria of the maximum value of the two sub-criteria is equal to 3.056452. Based on these data, overall it can be concluded that the usability level at Ok Google for the efficiency criteria is at a good level because the average reaches number 3.

3. Memorability

The total value for the memorability criteria is obtained from the sum of the results on the sub-criteria, the ease of the user remembering how to use Ok Google as a first timer user and ease on remembering how to use Ok Google when it has not been used for a long time. The average value for the sub-criteria for the ease of remembering on how to use Ok Google for the first time is 3.169355 from the maximum value of each sub-criteria is 4, the average for the sub-criteria of ease to use Ok Google when it has not been used for a long time is 3.032258 from the maximum value of each sub-criteria, and the average for the ease sub-criteria. So, the total average value for the memorability criteria of the maximum value of the two sub-criteria is equal to 3.100806. Based on these data, overall it can be concluded that the usability level at Ok Google for memorability criteria is at an easy-to-understand level because the average reaches number 3.

4. Errors

The total value for the errors criteria is obtained from the sum of the results in the sub-criteria, the user does not find errors or functions that are not working properly in the use of Ok Google and Ok Google provides the problem feature or help to solve the problem. The average value for the sub-criteria of the user not finding an error or the function running incorrectly in the use of Ok Google is 2.467742 from the maximum value of each sub-criteria is 4, the average for the Ok Google sub-criteria provides the help feature to overcome the problem that occurs is 2.846774 from the maximum value of each sub-criteria, and the average for the ease sub-criteria. So, the total average value for the criteria for errors of the maximum value of the two sub-criteria is 2.657258. Based on these data, overall it can be concluded that the usability level on Ok Google for errors criteria is at an elusive level because the average reaches number 2.

5. Satisfaction

The total value for the satisfaction criteria is obtained from the sum of the results in the sub-criteria, the users feel unencumbered (no need to take training) to access Ok Google and feel comfortable with the facilities available on Ok Google. The average value for the sub-criteria of the user unencumbered to access Ok Google is 3.129032 from the maximum value of each subcategory is 4, the average for sub-criteria users feel comfortable with the facilities available on Google OK is 2.975806 from the maximum value of each sub-criteria. So, the total average value for criteria errors from the maximum value of the two sub-criteria is equal to 3.052419. Based on these data, overall it can be concluded that the usability level on Ok Google for satisfaction criteria is at an easy-to-understand level because the average reaches number 3.

Furthermore, the average usability value obtained by Ok Google in Indonesia is 2.985484 or rounded up to 3.0000. If the average value is included in the predetermined category, then overall it can be seen that Ok Google is included in the category of usability "easy to understand" because the overall average reaches number 3.

After usability measurements on Ok Google, it was found that Ok Google has the highest memorability level and the lowest error. In addition, some of the findings obtained in the formulation of criteria. Many things are found in the assessment of Ok Google's usability level, one of them is almost all college websites do not pay attention to the convenience of users to find the help feature when facing difficulties in using Ok Google. To find out the usability of Ok Google, more details are needed on the respondent, and the comparison of usability ratings is done with methods other than Nielsen.

4 Conclusion

The results of the study show that Ok Google is quite capable of meeting user needs. It can be seen from learnability, efficiency, memorability, errors and satisfaction held by respondents. In addition, this study also illustrates the usability of Ok Google, which is the feature that currently exists and has been implemented must continue to be sought to experience a process of improvement and innovation. It seems that Ok Google is one of the innovations that are attractive to today's users. Technological developments and differences in generations that always innovate are one of the right reasons to justify. Millennials who are technologically literate will find it easier to absorb innovation using speech recognition that is applied appropriately and is able to accommodate learnability, efficiency, memorability, errors and user satisfaction in utilizing technology.

References

1. Qustoniah, A.: Perancangan Sistem Pengenalan Pola Suara Menggunakan Metode Piecewise Dynamic Time Warping. Widya Tek. **20**(1), 11–17 (2012)
2. Ronando, E., Irawan, M.I.: Pengenalan Ucapan Kata Sebagai Pengendali Metode Linear Predictive Coding—Neuro Fuzzy. J. Sains dan Seni ITS **1**(1), 51–56 (2012)
3. Jaya, M.T., Puspitaningrum, D., Susilo, B.: Penerapan Speech Recognition Pada Permainan Teka-Teki Silang Menggunakan Metode Hidden Markov Model (HMM) Berbasis Desktop. J. Rekrusif **4**(1), 119–129 (2016)
4. S. Google: Use 'Ok Google' Voice Search and Actions [Online]. Available: https://support.google.com/websearch/answer/2940021?co=GENIE.Platform%3DAndroid&hl=en. Accessed 08 Dec 2018
5. Kusumastuti, D.L., Yossy, E.H., Rabiha, S.G., Indrianti, Y.: Usability efficiency analysis on e-learning websites. **7**, 215–218 (2018)
6. N. J, Usability 101: Introduction to Usability. Nielsen Norman Group (2012)
7. ISO 9241-11:1998, Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs)—Part 11: Guidance on Usability

8. I. User Interface Engineering. www.usability.gov (2004)
9. U.S. Department of Health and Human Services, Usability Basics, U.S. Government. http://www.usability.gov/. Retrieved 27 Sept 2018
10. Creswell, J.W.: Research Design: Qualitative, Quantitative and Mixed Methods Approaches, 4th edn, p. 2014. Sage, London (2014)
11. Spiegel, M.R., Stephens, L.J.: Schaum's Outlines® Statistics, 6th edn. M. G. H. Education, New York (2018)
12. McDonald, M., Musson, R., Smit, R.: The Practical Guide to Defect Prevention. Microsoft Press, Washington (2008)
13. Turhan B, Bener AB (2007) Software Defect Prediction: Heuristics for Weighted Naive Bayes. In: Icsoft, vol. SE, pp. 244–249
14. Vallon, R., da Silva Estácio, B.J., Prikladnicki, R., Grechenig, T.: Systematic literature review on agile practices in global software development. Inf. Softw. Technol. **96**, 161–180 (2018)
15. Triola MF (2015) Essentials of Statistics, 5th edn, vol. 53, no. 9. Boston, Pearson
16. Rosas, S.R., Ridings, J.W.: The use of concept mapping in measurement development and evaluation: application and future directions. Eval. Program Plann. **60**, 265–276 (2017)
17. Kilic, S.: Cronbach's alpha reliability coefficient. J. Mood Disord. **6**(1), 47 (2016)

Part II
Machine Learning and Decision Support Systems

The Data Dimensionality Reduction and Features Weighting in the Classification Process Using Forest Optimization Algorithm

Daniel Kostrzewa and Robert Brzeski

Abstract The paper presents the data dimensionality reduction in the classification process, with a special presentation of using the ability of features weighting by determining the level of importance of a given attribute in the data vector. This reduction was implemented using the Forest Optimization Algorithm (FOA) and the use of a classifier allowing to enter the importance of each attribute for a data vector. The paper presents both, a description of the capability of using the FOA algorithm as well as the possibility of introducing modifications which allows to regulate the objective function between the obtained classification result and the number of reduced features. The conducted tests and obtained results were also presented. At the end of paper, a summary and the final conclusions are provided.

Keywords Dimensionality reduction · Feature selection · Features weighting · Classification · Weka · UCI machine learning repository · Sound · Forest optimization algorithm

1 Introduction

One of the stages of the classification process is proper preparation of the data set, including the data selection process. A typical action in data selection is to exclude a given feature, in a situation where it has no impact on the quality of the classification, or the impact is negative (i.e., the classification result deteriorates).

The process of removing unneeded data (features from the vector) has sometimes very strong impact on reducing the computational complexity of the classification process. When collecting data, there are often no guidelines about how each attribute

D. Kostrzewa (✉) · R. Brzeski
Institute of Informatics, Silesian University of Technology, Akademicka 16, 44-100 Gliwice, Poland
e-mail: daniel.kostrzewa@polsl.pl

R. Brzeski
e-mail: robert.brzeski@polsl.pl

M. Huk et al. (eds.), *Intelligent Information and Database Systems: Recent Developments*, Studies in Computational Intelligence 830,
https://doi.org/10.1007/978-3-030-14132-5_8

affects the possibility of correct classification. As a result, the number of gathered features is as high as possible—hoping that some of them will have a positive effect on the final result. At the same time, many of these features are only an unnecessary noise, or even have a negative impact on the classification process. Therefore, initially we usually collect the largest possible set of attributes. On the other hand, due to computational complexity and data denoising, there is a need to keep in vectors only those features, that positively affects the quality of classification. The selection process leaves these features but typically also those that have negligible or minor impact. In a standard selection process the level of influence of a given feature on the classification result is not determined. The influence is discrete: 0 or 1—so-called: exclude or include the attribute in a data vector.

On the other hand, it is possible to determine the importance of a given feature on the result of classification in a continuous interval [0, 1] and then use this information by the appropriate classifier. The word 'appropriate' is very important here. Namely, it must be a classifier that will be able to use information about the weight, the degree of importance of a given feature. Such a classifier is, for example the KNN.

The process of determining the degree of relevance of individual features can be achieved with the usage of Forest Optimization Algorithm (FOA).

1.1 Contribution

In this paper we will try to answer the following question. Is it possible to combine the features weighting (to keep the result of classification process at the highest rate) with the feature reduction (in order to reduce the data dimension)? To investigate this, several improvements to the FOA have been proposed.

1.2 Paper Structure

Appropriate related work in the subject is provided in Sect. 2. In Sect. 3 the standard version of FOA is described altogether with the proposed improvements. The conducted experiment results have been included in the Sect. 4. Summary and the final conclusions are provided in the Sect. 5.

2 Related Work

The data dimensionality reduction [10, 15] can be done by two types of methods. Attribute selection [7, 16] is based on the choice of a subset of features from a full set in such a way as to obtain the highest possible quality of classification [8]. Both, deterministic [13] or non-deterministic methods [18, 25, 26] are able to achieve good results. The second group of data dimensionality reduction strategies is construction of attributes [11] which produces new set of features as linear or non-linear combination of the original ones.

There are many non-deterministic methods of attribute selection. The FOA [9, 10], used in the current research, is one of the newest strategies which gave very good results on the field.

The FOA has also the ability to determine the degree of importance of features, its degree of influence on the classification result. Relevant research in this area is presented in papers by the authors of FOA [9, 10]. It shows not only the possibility of determining the weights of features [3, 5, 17, 21] but also the use of this knowledge in the classification process, using the KNN classifier [2], giving the ability to determine the weight of individual feature for the input data. The issue is not trivial, because most of the classifiers [24] cannot use knowledge of the features weight. All attributes are treated with the same relevance.

If FOA is able to determine the relevance of the feature and, on the other hand, has the possibility to select a subset of attributes from the whole set, then it should be achievable to combine this two objectives in one optimization process.

The obtained classification results can be compared according to many indicators [1, 4, 6, 19]. Choosing the right one is usually associated with the goal of classification. In this paper, as well as in the papers about FOA [9, 10], the accuracy was used to determine and compare the quality of the classification process.

3 Overview of Proposed Method

During the study the Forest Optimization Algorithm was used, in order to perform the dimensionality reduction of data vectors, as well as the process of determining the importance level of the features.

3.1 Forest Optimization Algorithm

In this work, in order to be able to carry out research, the FOA has been implemented in Java, at the beginning according to the original description [9, 10] (Fig. 1), together with adaptation to the problem being solved.

Each tree in FOA represents a potential solution of the problem. In this case it is an ordered sequence of numbers from the range [0, 1]. The numbers denote the relevance of the attributes.

In step 1 the initialization of given number of trees is performed. This phase consists of a random selection of a subset of features from the whole set. The age of new trees is set to 0.

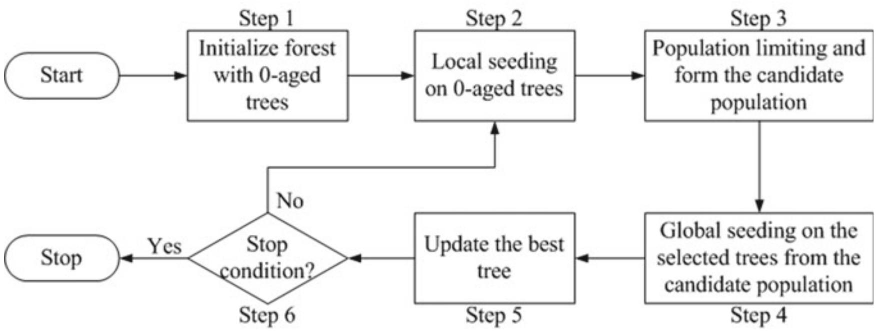


Fig. 1 The flowchart of FOA [10]

Local seeding is done afterwards (step 2). It is done only on trees with age equals 0. The seeding is equivalent to a random change of value of selected feature from the given range (in our case ± 0.2), with simultaneous limitation of the final result to the range [0, 1]. The local seeding counter (LSC) is a parameter which determines the number of newly created trees. After this step the age of all trees is incremented (except new solutions).

Step 3 is limiting the current population to a given number (area limit) by removing the oldest and worst performing results, and transferring them to the candidates population.

Global seeding (step 4) is performed only on each tree from the candidates population. It contains the random change (in range [0, 1]) of GSC number of features on every potential solution, where GSC (global seeding counter) is a parameter of the method.

If any of the new trees represents better solution than current best one, then in step 5 the best tree is saved as potential solution of the whole method.

The stop condition is checked in step 6. It may be the execution of a certain number of iterations, achievement of the expected accuracy of the results or lack of the improvement of the results in a given number of subsequent iterations. In the conducted research, the algorithm stop criterion is the lack of finding a better solution from currently the best, through 300 iterations.

3.2 Proposed Improvements

If the change of weights in a given range is replaced by setting the relevance to 0 or 1 (where 0 means attribute removal), a reduction of dimensionality is obtained. On the other hand, there is no information about the relevance of individual features.

Therefore, in this paper we propose and test the use of FOA, and its modifications to obtain both: the reduction of features and the information on the relevance of individual attributes (weight of features).

For this purpose, the following have been introduced by the authors:

- **Epsilon**—a parameter which allows to reduce the feature when the result (e.g., accuracy) is worse by no more than this parameter comparing to the one without reduction of the feature. Thus, theoretically there is a small possibility of deterioration of the classification accuracy by this value at each iteration in case of rejecting a given features. However, this situation in unlikely to happen, because sometimes there is an improvement of the classification result after reduction of some features. Secondly it affects only the result for the training and validation data. The final quality is checked on the test data—this is the reason why the final value can be quite different. In other words, epsilon parameter can be treated as method which allows to leave local minimum of the objective function. The tests were carried out for the values equals 0.001 and 0.0025.
- **Threshold** for removing features—if after the iteration of the algorithm, the weight of the feature will be reduced below this value, it is removed from the data vector—as very little relevant. This improvement enables faster removal of attributes of minor importance from the classification quality perspective. Current tests were carried out for the value equals 0.1.

3.3 Classification with Weighted Features

In the tasks of recognizing the belonging of an object to a particular class, the classifiers are used as a basic tool. They allow, based on a set of training objects, to classify the examined object into a given class. In this work the KNN algorithm was used due to the possibility of entering weights of the attributes.

The most popular method of determining the similarity of objects, in the KNN classifier, is to determine the value of the Euclidean distance between them. The function of the standard version of the distance is as follows (formula 1):

$$Euclidean\ distance(X, Y) = \sqrt{\sum_{n=1}^{D}(x_i - y_i)^2}, \tag{1}$$

where x_i and y_i are values from X and Y data vectors, between which the distance is measured, and D is the size of the data vector.

In this work, in order to optimize the efficiency of classification, the method of assigning weights to each of the features was used. This assumption introduces a modification to the formula for the Euclidean distance as follows [9] (formula 2):

$$Weighted\ Euclidean\ distance(X, Y) = \sqrt{\sum_{n=1}^{D} w_i (x_i - y_i)^2}. \tag{2}$$

The variable w_i determines the weight of the i-th feature. The modification allows the use of the KNN method to perform the classification process, using knowledge about the importance of individual features.

4 Results of the Conducted Experiments

The research was carried out on two datasets. First one, Dermatology [23] (34 features, 366 vectors, 6 classes), is well known dataset in the literature taken from UCI Machine Learning Repository. The second dataset (AudioFeelDB [14], 82 features, 200 vectors, 2 classes) concerns the classification of songs by its genre. This is an important problem [20, 22], which is why this dataset has been taken into account.

The datasets have been divided into 10 parts in such a way that the number of representatives/vectors of a given class in each part is the same (or differs by at most 1 vector). Then, from these 10 parts, 10 different training and test sets were created at ratio of 70/30 (i.e., 70% of vectors are in training set, and remaining 30%—in test set). The training group was used to train the classifier according to FOA outcome using internally the 10-fold cross-validation method.

The experiments were carried out as follows: on each of the 10 sets (training set with test counterpart), three FOA optimizations were performed. As a result, a total of 30 tests were done for the both datasets.

The test software we made in Java and we exploited the WEKA library [12]. All computations were performed on the machine with AMD FX-8350 Eight Core Processor 4.00 GHz with 32 GB of DDR3 RAM.

In order to check the objectives of the paper, we designed a set of 6 different tests:

- **Test 1** This test is carried out without determining the weights of individual features. The main goal is to obtain the largest reduction of attributes, assuming that the result of classification does not deteriorate from the initial value for the full set of features (without reduction). The FOA algorithm is used here with a discrete change of weights: 0 or 1 (i.e., the attribute is excluded or included in the subset respectively). So the goal is to remove as many features as possible without worsening the result of the classification.
- **Test 2** The test determines the weights of individual features. This is done without focusing on reducing the quantity of features. The aim is to obtain the best classification result for the full data vector (but with specified weights of individual features). This is the test where at the same time the weights of some features have been reduced to value 0. This is a reason why some of attributes have been removed. The FOA algorithm is used here in the way, where weight of features can be change by a random value from the given range (in our case $\pm$ 0.2), with simultaneous limitation of the final result to the range [0, 1].
- **Test 3** The test is the same as in Test 2 but with exploitation of the epsilon parameter. The value of epsilon is equal 0.001.

- **Test 4** This is an extension of Test 3. We used the threshold for removing features with value equal 0.1 (i.e., when the weight of given attribute is smaller than 0.1 the weight is set to 0 in another solution).
- **Test 5** This situation is a copy of Test 3 but the value of epsilon is equal 0.0025.
- **Test 6** The test is the same as the Test 4 but the epsilon parameter is equal 0.0025.

To summarize designed tests it can be said that Test 1 and Test 2 are the baseline and can be treated as two versions of standard FOA. Tests 3–6 show results for proposed improvements.

The results of the tests are gathered in Tables 1 and 2 for Dermatology and AudioFeelDB datasets respectively.

In Tables 1 and 2 the column *Acc* is the average accuracy value of classification process for the test data after 30 runs of FOA algorithm. The column *Qty* is the average quantity of removed attributes through the FOA algorithm. The rows *Change % to Test* is the percentage value of the change that was made comparing to a results from particular test.

Test 1 shows the possibilities of reduction of features and the quality of classification for such level of reduction. The values obtained in this test is the baseline for the rest of tests.

Test 2 illustrates the quality of the classification which was obtained by using the feature weighting, and how many weights were zeroed (how many attributes were removed from a subset) during the FOA algorithm optimization. As can be seen it was possible to improve the accuracy (comparing to test 1, by 3.4% for Dermatology and 4.6% for AudioFeelDB datasets) at the expense of a significant reduction in the number of features removed.

The results of Tests 3 and 5 shows that the use of the feature weighting together with introducing the epsilon parameter with values 0.001 and 0.0025 in the FOA algorithm considerably increases the number of removed features from the final subset (by over 200% in comparison to Test 2). The quality of classification drops by less than 1% for Dermatology dataset, and was increased by 3.3 and 6% for AudioFeelDB dataset for Tests 3 and 5 respectively.

Tests 4 and 6 prove that further removal of attributes is possible by incorporating reduction threshold parameter with value 0.1 in the FOA algorithm without significant change of the classification quality. Comparing to test 2 FOA removed over 250% more features. Moreover, the accuracy is even higher in three of four cases. The results are better than outcomes which were obtained in Tests 3 and 5.

Epsilon parameter by assumption gives the possibility to further reduce number of features while accepting the worsening of the classification result (at least initially—during the learning process). In order for this deterioration to be sufficiently small or did not occur at all, the epsilon value have to be relatively small. Therefore, for the conducted tests, such low values were adopted. In the case that the emphasis would be more focused on the larger reduction of the number of features than on the non-deterioration of the results, the epsilon values could be set to higher values.

The analysis of the results gathered in Tables 1 and 2 shows that the Epsilon parameter (Test 3 and 5) fulfilled its role even better than the initial assumptions. It

Table 1 Results of experiments for Dermatology dataset

	Test 1		Test 2		Test 3		Test 4		Test 5		Test 6	
	Acc	Qty	Acc	Qty	Acc	Qty	Acc	Qty	Acc	Qty	Acc	Qty
Average of 30 runs	0.918	**18.6**	**0.949**	2.5	0.942	7.7	0.941	8.8	0.942	8.3	**0.948**	9.0
Standard deviation	0.029	2.32	0.017	1.38	0.020	1.77	0.019	1.82	0.017	1.44	0.018	1.28
Change % to Test 1	–	–	3.4	−86.6	2.6	−58.6	2.5	−52.7	2.6	−55.4	3.8	−51.6
Change % to Test 2	–	–	–	–	−0.74	208	−0.84	252	−0.74	232	0.1	360
Change % to Test 3	–	–	–	–	–	–	−0.1	14.3	0	7.8	0.6	16.9
Change % to Test 4	–	–	–	–	–	–	–	–	0.10	−5.7	0.6	2.3
Change % to Test 5	–	–	–	–	–	–	–	–	–	–	0.6	8.4

Table 2 Results of experiments for AudioFeelDB dataset

	Test 1		Test 2		Test 3		Test 4		Test 5		Test 6	
	Acc	Qty	Acc	Qty	Acc	Qty	Acc	Qty	Acc	Qty	Acc	Qty
Average of 30 runs	0.613	**59.7**	0.641	8.1	**0.662**	19.3	0.652	21.6	0.650	21.4	0.651	22.6
Standard deviation	0.052	9.88	0.068	2.49	0.061	3.26	0.062	2.54	0.063	4.26	0.071	2.87
Change % to Test 1	–	–	4.6	−86.4	8.0	−67.7	6.4	−63.8	6.0	−64.2	6.2	−62.1
Change % to Test 2	–	–	–	–	3.3	238	1.7	267	6.0	264	1.6	279
Change % to Test 3	–	–	–	–	–	–	−1.5	11.9	−1.8	10.9	−1.7	17.1
Change % to Test 4	–	–	–	–	–	–	–	–	−0.3	−0.9	−0.2	4.6
Change % to Test 5	–	–	–	–	–	–	–	–	–	–	0.2	5.6

turned out that in comparison to Test 2 (without using the Epsilon parameter), not only the number of removed features has increased but also for the AudioFeelDB dataset the quality of classification has improved.

Threshold—in the range of the possible weight value of the feature (the range [0, 1]—where weight 0 means no impact of feature on the classification result) the authors determined arbitrarily that the weight in the range up to 0.1 reflects the state, of so little influence, that it maybe is worth to remove that feature.

For the Threshold parameter the analysis of results (Tables 1 and 2) shows that, like the epsilon parameter, it also fulfilled its role even better, than the initial assumptions. It turned out (comparing Test 4 to Test 3 and Test 6 to Test 5) that not only number of removed features has increased, but also for Test 6, the quality of classification has improved.

Generally, it can be seen that the choice of parameter value is not a trivial issue and relates to multi-parameter optimization (quality of classification and the number of removed features depending on the parameters Epsilon and Threshold). The values of these parameters can be appropriately selected depending on the objective function.

In summary, introducing the new improvements in FOA, enabled new path of research which combine two, opposite at the first glance, objectives.

5 Conclusions

Using the optimization algorithm (FOA in case of this research) does not force to choose one of the options: improvement of classification results or reducing dimensionality. It turns out that with the appropriate modifications of original FOA strategy, it is possible to combine these goals. At the same time, depending on the needs, the final result can be adjusted between a greater dimensionality reduction and improvement of the quality of classification process. It can be done through adequate setting of the value of the epsilon parameter and the reduction threshold. Therefore, by analyzing provided results, it can be said that for this datasets, the proposed modifications gives positive results.

With introduced improvements the additional information about the degree of importance of a given feature is discover at the same time as reducing the dimensionality. It is important that we can obtain a certain domain knowledge from a given area—information which feature have the influence on the final result.

We have to admit that the conducted research is not only the continuation of our previous work [13, 14] but also the introduction to the future one. Our ongoing research focuses on the determination of degradation of the classification quality while increasing the number of removed features for many test data. We are planning to use other methods on this field, e.g., reinforcement learning for attribute selection and deep artificial neural networks for performing classification process.

Acknowledgements This work was supported by BKM-509/RAU2/2017 (DK) and BK-213/RAu2/2018 (RB) grants from the Institute of Informatics, Silesian University of Technology, Gliwice, Poland.

References

1. Agrawal, R., Imielinski, T., Swami, A.: Database mining: a performance perspective. IEEE Trans. Knowl. Data Eng. **5**(6), 914–925 (1993)
2. Aha, D.W., Kibler, D., Albert, M.K.: Instance-based learning algorithms. Mach. Learn. **6**(1), 37–66 (1991)
3. Aha, D.W.: Feature weighting for lazy learning algorithms. In: Feature Extraction, Construction and Selection, pp. 13–32. Springer (1998)
4. Arie, B.D.: Comparison of classification accuracy using cohen's weighted kappa. Expert. Syst. Appl. **34**(2), 825–832 (2008)
5. Bach, M., Werner, A.: Cost-sensitive feature selection for class imbalance problem. In: International Conference on Information Systems Architecture and Technology, pp. 182–194. Springer (2017)
6. Costa, E., Lorena, A., Carvalho, A., Freitas, A.: A review of performance evaluation measures for hierarchical classifiers. In: Evaluation Methods for Machine Learning II: Papers from the AAAI-2007 Workshop, pp. 1–6 (2007)
7. Doak, J.: An evaluation of feature selection methods and their application to computer security. Technical Report, Department of Computer Science, University of California, Davis CA (1992)
8. Gasca, E., Sánchez, J.S., Alonso, R.: Eliminating redundancy and irrelevance using a new mlp-based feature selection method. Pattern Recognit. **39**(2), 313–315 (2006)
9. Ghaemi, M., Feizi-Derakhshi, M.R.: Forest optimization algorithm. Expert. Syst. Appl. **41**(15), 6676–6687 (2014)
10. Ghaemi, M., Feizi-Derakhshi, M.R.: Feature selection using forest optimization algorithm. Pattern Recognit. **60**, 121–129 (2016)
11. Guyon, I., Elisseeff, A.: An introduction to variable and feature selection. J. Mach. Learn. Res. **3**(Mar), 1157–1182 (2003)
12. Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I.H.: The weka data mining software: an update. ACM SIGKDD Explor. Newsl. **11**(1), 10–18 (2009)
13. Kostrzewa, D., Brzeski, R.: The data dimensionality reduction in the classification process through greedy backward feature elimination. In: International Conference on Man–Machine Interactions, pp. 397–407. Springer (2017)
14. Kostrzewa, D., Brzeski, R., Kubanski, M.: The classification of music by the genre using the knn classifier. In: International Conference: Beyond Databases, Architectures and Structures, pp. 233–242. Springer (2018)
15. Liu, H., Motoda, H.: Feature Extraction, Construction and Selection: A Data Mining Perspective, vol. 453. Springer (1998)
16. Liu, H., Motoda, H.: Feature Selection for Knowledge Discovery and Data Mining, vol. 454. Springer (2012)
17. Özşen, S., Güneş, S.: Attribute weighting via genetic algorithms for attribute weighted artificial immune system (awais) and its application to heart disease and liver disorders problems. Expert. Syst. Appl. **36**(1), 386–392 (2009)
18. Papakostas, G., Koulouriotis, D., Polydoros, A., Tourassis, V.: Evolutionary feature subset selection for pattern recognition applications. In: Evolutionary Algorithms. InTech (2011)
19. Powers, D.M.: Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation (2011)
20. Rosner, A., Weninger, F., Schuller, B., Michalak, M., Kostek, B.: Influence of low-level features extracted from rhythmic and harmonic sections on music genre classification. In: Man-Machine Interactions, vol. 3, pp. 467–473. Springer (2014)

21. Tosun, A., Turhan, B., Bener, A.B.: Feature weighting heuristics for analogy-based effort estimation models. Expert. Syst. Appl. **36**(7), 10325–10333 (2009)
22. Tzanetakis, G., Cook, P.: Musical genre classification of audio signals. IEEE Trans. Speech Audio Process. **10**(5), 293–302 (2002)
23. UCI Machine Learning Repository: Dermatology data set. Irvine, CA: University of California, School of Information and Computer Science. http://archive.ics.uci.edu/ml/datasets/dermatology
24. Wu, X., Kumar, V., Quinlan, J.R., Ghosh, J., Yang, Q., Motoda, H., McLachlan, G.J., Ng, A., Liu, B., Philip, S.Y., et al.: Top 10 algorithms in data mining. Knowl. Inf. Syst. **14**(1), 1–37 (2008)
25. Yang, J., Honavar, V.: Feature subset selection using a genetic algorithm. In: Feature Extraction, Construction and Selection, pp. 117–136. Springer (1998)
26. Yang, Z.M., He, J.Y., Shao, Y.H.: Feature selection based on linear twin support vector machines. Procedia Comput. Sci. **17**, 1039–1046 (2013)

Recent Developments on Evolutionary Computation Techniques to Feature Construction

Idheba Mohamad Ali O. Swesi and Azuraliza Abu Bakar

Abstract The quality of the search space is an important factor that influences the performance of any machine learning algorithm including its classification. The attributes that define the search space can be poorly understood or inadequate, thereby making it difficult to discover high quality knowledge and understanding. Feature construction (FC) and feature selection (FS) are two pre-processing steps that can be used to improve the feature space quality, by enhancing the classifier performance in terms of accuracy, complexity, speed and interpretability. While FS aims to choose a set of informative features for improving the performance, FC can enhance the classification performance by evolving new features out of the original ones. The evolved features are expected to have more predictive value than the originals that make them up. Over the past few decades, several evolutionary computation (EC) methods have been proposed in the area of FC. This paper gives an overview of the literature on EC for FC. Here, we focus mainly on filter, wrapper and embedded methods, in which the contributions of these different methods are identified. Furthermore, some open challenges and current issues are also discussed in order to identify promising areas for future research.

Keywords Classification · Evolutionary computation · Feature selection · Feature construction

1 Introduction

In machine learning and data mining, classification is a crucial task that targets to classify an unseen observation to its corresponding category [1]. However, in

I. M. A. O. Swesi (✉) · A. A. Bakar
Faculty of Information Science and Technology, Centre for Artificial Intelligence Technology, University Kebangsaan Malaysia, Bangi, Selangor, Malaysia
e-mail: idheba@siswa.ukm.edu.my

A. A. Bakar
e-mail: azuraliza@ukm.edu.my

M. Huk et al. (eds.), *Intelligent Information and Database Systems: Recent Developments*, Studies in Computational Intelligence 830,
https://doi.org/10.1007/978-3-030-14132-5_9

many real-world problems, several learning algorithms may not accomplish good classification performance. One possible reason for this is due to the poor quality of the search space which is represented by a set of attributes that has been grown enormously. While earlier FS studies have used less than 100 features [2], the current literature can deal with thousands if not millions of features which are not useful for classification purposes as they usually contain many irrelevant and redundant features [3]. The presence of these features not only results in decrease of the classification accuracy but also increase the computational cost [4]. Thus, the quality of search space can be improved through FS and FC by decreasing the dimensionality, speeding up the learning process and enhancing algorithm's performance [5–7]. However, even though both tasks have been studied for decades, challenges continue to persist because of the large search space. For a given dataset with N features, the FS method has to consider 2^N possible subsets to obtain good solutions. In contrast, FC method needs to search for good features, select appropriate operators and choose a good way to apply these operators on the original features. According to [5], FC is defined as a process of discovering useful knowledge regarding hidden relationships among features in data where the new features are either used alone or by augmenting the original data.

FC approaches can be categorised into three types: filter, wrapper and embedded [8, 9]. In filters, the features are constructed based on the intrinsic characteristics of the data, whereas the wrappers use a learning algorithm to assess the constructed features. However, few demerits are associated with the wrapper FC approaches, including high computational cost and losing generality, but better discriminating ability is also rendered compared to filters. Unlike wrappers and filters, embedded approaches perform FC and simultaneously build a classifier.

FC is typically an NP-hard problem due to the huge search space that has increased exponentially alongside the total number of features and the function operators. Thus, this makes FC methods often being stuck into local optima and computationally intensive [10]. Besides the large search space, the FC faces challenges due to the complex interactions amongst features [11]. Therefore, a global search technique should be used in designing FC methods to construct high-quality features. Recently, EC techniques have been widely employed for classification due to their global search ability. These techniques including genetic programming (GP), genetic algorithm (GA), and particle swarm optimisation (PSO) which have demonstrated their success to many areas and among them is feature construction [10, 12–15]. Although a significant level of work has been undertaken in this field, there is no comprehensive guideline to identify the weaknesses and strengths of these approaches. This can introduce new opportunities for the researches to conduct improved and successful works. Also, while there have been many papers put forward over the last five years, no review papers had been published since 2009 [16]. Therefore, this paper aims to provide an overview of past studies that are conducted for FC problems using EC approaches.

2 EC for Feature Construction

Evolutionary computation (EC) is a group of global optimization techniques that were inspired by the principles of biological evolution. Generally, EC composed of evolutionary algorithms (EA) and swarm intelligence (SI), which have gained a huge attention and obtained good results. Although these approaches are widely used for FS, only three of them are employed for FC namely; GA, GP and PSO. However, compared with GA and PSO, GP is the most utilised algorithm for FC because of its tree-based representation that has the positional to deal with function operations easily and directly.

Figure 1 presents the overall categories of EC for FC methods. These methods can be classified into various categories based on the three components: the EC branches, the evaluation and the number of objectives. For the two EC branches, only three techniques have been used for FC and they are discussed in this paper. The GA and GP are two examples of EA, and PSO is an example of SI. Depending on the evaluation criteria, the review studies will be based on three approaches: filter, wrapper and embedded. Since there are two conflicting objectives in FC which are increasing the classification accuracy and reducing the dimensionality, the current EC based FC approaches can be classified into single objective and multi-objective approaches. The following subsections summarise and discuss the recent studies conducted on the three approaches for FC as can be seen in Table 1.

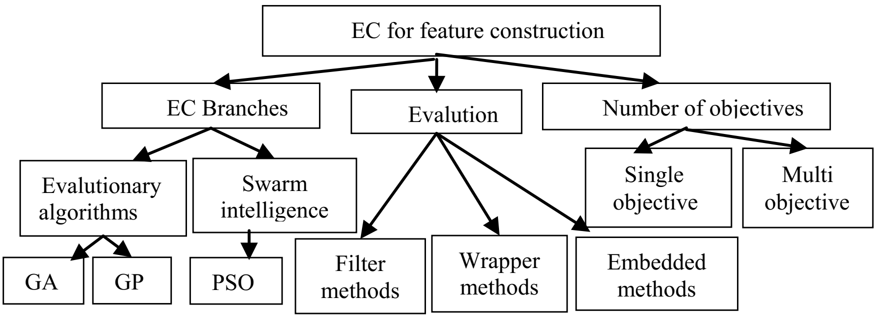


Fig. 1 Overall categories of EC for feature construction

Table 1 Taxonomy of the reviewed EC based FC approaches

Category	Single constructed feature	Multiple constructed features
Wrapper	[10, 42, 43]	[13, 17–19, 26, 28, 30]
Filter	[33, 38]	[20–23, 37]
Embedded	[39]	[24, 25, 39, 40]

2.1 Genetic Algorithm

Since 1995, several GA based FC methods were introduced, where the work of Vafaie and De Jong [13] was among the earliest studies that investigated its use for FC. In this study, the GA is combined with GP for creating multiple features and C4.5 was utilised to evaluate the feature subset. Experiments on one texture dataset show slightly improved performance. However, only one dataset with eight features was used, which is not sufficient to verify the effectiveness of the proposed method.

For wrapper approaches, different methods have been introduced. For example, Lillywhite et al. [17] proposed an evolution-constructed (ECO) method for general object recognition. In this method, a standard GA is employed to automatically construct the features and identify highly discriminative features. The result of ECO method was found to be better or comparable versus other methods, including partial least squares (PLS) and convolutional restricted boltzmann machine (C-RBM). In other works such as [18], an evolutionary method was used for both FS and FC, in which the process of creating multiple features was optimised by GA. Presented results show that the proposed method gave good results. However, the method did not perform well on the entire set of features thus FS is a must for reducing the dimensionality. In contrast to [17, 18] that utilises an algebraic representation (set of operators) for FC, Shafti and Pérez [19] introduced another notion named non-algebraic representation to reduce the difficulty of constructing features and improve the performance. The proposed method called MFE3/GA^{DR}, where GA helps to deal with the problem of interactions better than other FC methods. Results revealed that the MFE3/GA^{DR} could construct highly informative features with better accuracy. However, the method was sensitive to noise data when missing or incorrect feature values are available.

For filter approaches, various measures have been employed with GA for FC. The authors in [20] applied GA to construct X-of-N multiple features where the information gain ratio (IGR) was used as fitness function. Then, the constructed features are combined with the original features before they send to C4.5 for the evaluation. Results on seven datasets ranging from 7 to 57 features showed that the performance of C4.5 was improved. Next, García et al. [21] utilised a set of function operators over the initial features to construct new ones. The relevance of the new constructed features was assessed through an information measure. Experimental study was carried out using 9 UCI datasets and the results show improved performance over GP, Chi and Ishibushi. However, due to the complexity of GA, only three operators were used while using more functions could give better results regarding the accuracy and efficiency.

Later, García et al. [22] proposed an improved approach by combining two different methods. One of the methods extracts relations between the initial attributes and the other one creates new attributes by employing functions over the initial attributes. An evaluation of this approach, the authors conducted their experiments on 27 UCI datasets and the results exhibited a good trade-off among the accuracy, interpretability and the time needed to obtain the model. In [23], another GA based filter approach

Table 2 Summary of GA based FC methods performed by different researchers

FC method	Studies	Advantages	Disadvantages	Assessment/dataset
Wrapper	Vafaie and De Jong [13]	Improved performance on the texture classification problem	Tested on one small dataset, which is not enough to verify the performance of the method	Train-test split/image data
	Lillywhite and Tippetts [17]	ECO performs better or comparable to other methods	Produced many features leading to reduced accuracy and increased computational time	Cross validation/image data
	Drozd and Kwasnicka [18]	The method provided higher accuracy than FS method	It does not perform well on the entire features thus FS is a must, tested on low features dimension	Train-test split/UCI data
Filter	García et al. [21]	The system gives more features and increased the accuracy	It used only three function operators while utilising more functions can improve the accuracy and efficiency	Use statistical test/UCI data
	García et al. [22]	The results showed a good trade-off among accuracy, interpretability and time	Tested on low features dimension. Except for the accuracy, metrics like specificity, recall, precision could be considered	Use statistical test/UCI data
	Alfred [23]	The proposed method increased the accuracy of the J48 learner	It was difficult to summarize the data for high feature number due to the increase of unique values	Cross validation/PKDD

has been discussed, in which the GA evolves new features out of the original features by considering four measures. (i.e., Davies-Bouldin index, information gain, information gain coupled with cluster entropy, total clusters entropy). Results show that the accuracy of the constructed features can be enhanced. Table 2 depicts a list of the aforementioned GA based FC methods with their advantages and disadvantages.

In summary, GA for FC started at least more than 20 years ago ([13] put forward in 1995). Since then, it has achieved a limited success on problems with hundreds of features, whereby most of them are wrapper methods. This resulted in a high computation time as GA mostly requires a large number of assessments, where every assessment in a wrapper method usually takes a long time. Further, the use of GA for FC in solving large scale problems has not been investigated, which requires a good representation that can lower the dimensionality of the feature space. Also, the design of genetic operators, e.g., crossover and mutation, will open opportunities for the researchers to combine the features and produce new ones.

2.2 Genetic Programming

Since GP owns a flexible tree in building programmers and functions, the algorithm has been widely utilised for FC, yielding competitive results compared to other EC methods. Different ways have been made for the use of GP in FC, for instance [24, 25], used a GP as an embedded approach whereas the authors in [26–32] applied GP as wrapper approach. GP as filter approach was used in [33–38]. One of the earliest works on GP for constructing multiple features was proposed in [26], where a C4.5 inducer was employed to evaluate the constructed features. Results showed that the proposed method achieved better performance regarding the accuracy and dimensionality. However, the method needs to specify the number of constructed features beforehand and it could also expose to over-fitting problem.

Another example of wrapper approaches was carried out by Ahmed et al. [28] who proposed two approaches called GPWFC1 and GPWFC2 for constructing multiple. A classification accuracy of random forest (RF) is served as fitness function of GPWFC1, whereas the p-value of ANOVA test and entropy gain of RF are employed as fitness of GPWFC2. Reported results display that the second method achieved better classification performance than the first method. However, as the two methods are wrappers based approaches, they have a potential limitation of high computation time. Next, a hierarchical feature construction approach for image classification was proposed by Suganuma et al. [30]. Experiments showed that the proposed method is superior compared to the existing FC methods.

For filter approaches, each individual is assessed using a filter measure. For instance, in [33], GP was used to construct single feature using four filter measures as fitness functions (i.e. GI, IG, chi-square and GI + IG). Experiments on five datasets ranges from 4 to 12 features showed that the constructed feature improved the performance of four classifiers (C5, CHAID, CART and MLP) without any bias for the four fitness functions applied in the FC process. In contrast to [33] that

Table 3 Summary of GP based FC methods performed by different researchers

FC method	Studies	Advantages	Disadvantages	Assessment/dataset
Wrapper	Krawiec [26]	Improved accuracy and reduced the dimensionality	Needs to specify the number of constructed features and it is exposed to overfitting	Use statistical test/UCI data
	Ahmed et al. [28]	Improved results in terms of accuracy and No. of features	Both methods have a high computation time and they are limited to binary classes	Use statistical test/MS data
	Suganuma et al. [30]	The method outperforms several existing FC methods	Tested on two datasets thus it needs to be evaluated on more challenging dataset	Cross validation/image data
Filter	Neshatian et al. [37]	Increasing the classification performance	Feature set being too big, tested on low dimensional features	Use statistical test/UCI data
	Liang et al. [38]	Improved accuracy over the original data	The single feature was not able to improve the accuracy	Train-test split/image
Embedded	Chen et al. [25]	Better regression result and reduced time	Constructed a large number of features, exposed to overfitting	Use statistical tests/UCI data
	Tran et al. [39]	Obtained competitive accuracy and reduced the dimensionality	It does not apply any technique to remove irrelevant and redundant features	Cross validation/microarray data
	Ahmed et al. [40]	Better results in both accuracy and feature set	Tends to produce slightly larger No. of constructed features	Train-test split/MS data

constructed only one feature, a GP based multiple feature construction (GPMFC) approach was proposed in [37]. Information-theoretic measure was used as fitness function to evaluate the constructed feature. Consistent enhancement in the accuracy was confirmed through the results from six datasets with 4–13 features. Next, Liang et al. [38], proposed a filter GP method (FGP), where a novel entropy based fitness function was used for constructing single feature. Results showed that the combined features (constructed + originals) perform better than the original features.

For embedded approaches, compared with GA, GP has the ability to perform embedded FC in which it can be used as both a search technique and a classifier [25, 39, 40]. Chen et al. [25] proposed a method called GP with embedded feature construction (GPEFC). Six high-dimensional datasets were used to assess the performance of the GPEFC and the results showed better performance over standard GP. Next, Tran et al. [39] proposed an embedded GP method for constructing both single and multiple features. Based on the results, the proposed method was able to significantly enhance the accuracy and reduce the dimensionality. However, since the datasets contain a huge number of redundant or irrelevant features, it is impractical to send all features for experimental validation as they may degrade the performance and increase the computational cost.

EC techniques, particularly GP algorithms are more suitable for multi-objectives optimization as they can produce multiple solutions rather than one solution [40, 41]. A GP based multi-objective FC (GPMOFC) approach was proposed in [40]. The main goal of this approach is to minimise the number of constructed features while at the same time maintaining high accuracy. Results showed that the GPMOFC achieved better performance than the multi-objective FS in most cases. A list of the abovementioned GP based FC methods and their advantages and disadvantages are summarized in Tables 3 and 4.

To sum up, GP has been successfully implemented for FC showing promising results with much work higher than other ECs. Compared with GA, GP has the potential to perform embedded FC where it can be used as a classifier. Besides, GP has shown promise in solving high dimensional data with thousands of features as it does not include all the features in the construction process. Even though the algorithm achieved good performance, it faced some limitations such as the high computational cost because it requires testing tens or hundreds of thousands of programs to find a desired solution. Another issue, in case of biological data where the number of features is large than the number of examples, GP may also exposed to overfitting.

2.3 Particle Swarm Optimisation

While an extensive research has been conducted to apply PSO for FS, relatively little research has been devoted to investigate its performance for FC with only three works [10, 42, 43]. The first PSO based wrapper FC (PSOFC) approach was presented in [42]. In this study, a single feature was constructed using the original features and function operators, where the last are selected via a local search. Experiments showed that the PSOFC approach could improve the classification performance on seven UCI datasets ranging (from 14 to 500) features. However, the local search employed in PSO for selection function operators is time consuming as it assesses all the operators in order to select the optimal ones. Later, Dai et al. [10] proposed two PSO based wrappers FC, (i.e. PSOFCArray and PSOFCPair) approaches. Results exhibited that the two methods were able to enhance the classification performance. Nonetheless, the PSOFCPair shared the same value that is useful in determining if the feature is chosen or not and which operator is selected, which may not be an ideal value for both feature and operator selection. Due to this limitation, the classification performance achieved by the PSOFCPair was degraded in most cases. More recently Mahanipour and Nezamabadi-pour [43] modified the two approaches in [10] by applying forward feature selection (FFS) method to reduce the dimensionality. Then the two modified approaches are used to construct the new features. Results showed promising results in increasing the classification performance. However, the experiments were only performed on datasets with small number of features ranging between 14 and 500. A list of the above PSO based FC methods and their advantages and disadvantages are summarized in Table 4.

Overall, the current studies have shown that PSO has a potential to address FC problems, but with some limitations. One of which is the high computational cost because the original representation of PSO does not have the ability to consider the nominal or categorical values (i.e. function operators). Thus, a new representation scheme is required to reduce the size of the search space, improve the computational time and help to understand the obtained results. The current PSO based FC approaches focus mainly on relatively small-scale problems with binary classes. Hence, there is an important need to further investigate the potential of PSO for FC and design more rigorous approaches that can handle large scale data with multiple classes accurately and efficiently.

Table 4 Summary of GP based FC methods performed by different researchers

FC method	Studies	Advantages	Disadvantages	Assessment/Dataset
Wrapper	Xue et al. [42]	Improved classification performance and reduced dimensionality	The method suffers from the problem of high computational cost	Train-test split/UCI data
	Dai et al. [10]	Improved performance in terms of accuracy and efficiency	PSOFCPair achieves a worse performance than other methods	Train-test split/UCI data
	Mahanipour and Nezamaba di-pour [43]	Better results in increasing the accuracy and reducing the dimensionality	Limited to datasets with small number of features	Train-test split/UCI data

2.4 Quantitative Analysis

In this section, we provide the results of various FC methods that were reviewed in this work using 19 UCI datasets with different number of examples (E) and features (F). Table 5 shows the comparison between these methods based on the classification accuracy (ACC) and the number of constructed features. As can be seen from the table, each method achieved consistent performance, where the GPMF method could obtained the best performance among the seven FC methods. Also, GP-IG, NSLV-F, NSLV-FR and PSOFC achieved good performance, while the PSOFC-Pair showed poorer performance but still showed consistent results by adding the constructed feature to the original features.

3 Current Issues and Challenges

Although recently there has been some success on FC area, many significant issues and challenges still faced and need to be addressed which are discussed as follows:

- Most of the FC approaches suffer from the problem of high computational cost, which is an important problem in EC techniques as they usually require a large number of evaluations. Although some filter methods are deemed faster when compared with wrappers, their performance is worse than wrappers [12]. Thus designing effective and efficient methods for FC problems is challenging. Some important aspects such as a fast evaluation criterion and an efficient search mechanism can be considered to speed up the process.
- Even though many studies have been carried out for FC, their application to large scale problems is still a challenge. Some EC methods have been put forward to deal with such problems [9, 14, 25, 28, 39, 44] but with some limitations, e.g. time complexity. Therefore, a novel search mechanism needs to be considered so as to explore the whole search space effectively and efficiently.
- The size of the search space is highly relies on the representation scheme, where the most popular one is a binary string of length equal to the total number of features. However, this representation restricted the performance of ECs on large scale problems and it also shows only if a feature is chosen or not which cannot be easily utilised for FC. Therefore, designing a good representation is an open issue, as it aims to overcome the large search space and improves the search ability.
- Generalisation is an important issue in EC for FC particularly when wrapper approaches are utilised. In this case, the constructed features may have the problem of overfitting to the training data which inevitably leads to a poor classification performance on unseen test data. Here, filter approaches are more general than wrapper approaches since they did not fit to any learning algorithm and their goal is to identify the patterns of the data.
- Memetic algorithms (MA) which are a combination of EC techniques with local search methods have shown a promising performance in FS by fine tuning the

Table 5 Comparison between various FC methods based on ACC/No. of constructed features

Dataset	#E	#F	NSLV-F [21]	NSLV-FR [22]	GPMFC [37]	GP-IG [33]	PSOFC [42]	PSOFC-Arry [10]	PSOFC-Pair [10]
Bupa	345	6	59.9/3	68.6/2	67.6/2	67.7/1	–	–	–
Bridges	106	11	57.5/1	58.5/4	–	–	–	–	–
Heart	303	13	78.4/1	78.1/2	–	–	–	–	–
Iris	150	4	94.6/3	94.6/2	95.7/3	–	–	–	–
Pima	768	8	74.8/1	75.7/1	–	–	–	–	–
Sonar	208	60	73.9/1	79.6/1	–	–	–	63.3/1	53.1/1
Tao	888	2	85.1/2	90.4/2	–	–	–	–	–
Blood	748	4	76/6	75.9/3	–	–	–	–	–
WBDC	569	30	95.6/1	95.2/2.5	–	–	93.3/1	91.5/1	86.5/1
WBDC	198	33	77.3/3	73.3/3	–	–	–	–	–
Balance	625	4	82.4/2	91.4/1	100/3	100/1	–	–	–
Wine	178	13	94.9/2	96.1/1	95.4/3	94.7/1	–	–	–
Thyroid	215	5	93.4/2	94.8/1	96.7/3	–	–	–	–
Australian	690	14	–	–	–	–	85.4/1	85.1/1	65/1
Ionosphere	351	34	–	–	–	–	81.1/1	76.7/1	75.5/1
Hillvalley	606	100	–	–	–	–	99.4/1	96.9/1	54.6/1
Musk1	476	166	–	–	–	–	65.9/1	58.8/1	58.9/1
Madelon	4400	500	–	–	–	–	53.9/1	57.3/1	50.9/1
Avg.			80.29	82.48	91.08	87.47	79.83	75.67	63.5

solutions generated by EC methods. However, the potential of MA in FC has not been explored to date.

- Some of the EC based FC methods focused on constructing a single feature rather than multiple features. Single feature might not obtain good classification accuracy, and also it might not sufficient for complex problems whereas multiple features are shown to be more promising, however, constructing multiple features has not extensively investigated.

4 Conclusions

This paper presented a review of work undertaken in the area of FC covering the most commonly EC methods are used in solving FC problems. The Current issues and challenges as well as some possible research directions have also discussed. The review shows that the EC methods have gained more attention in addressing FC tasks and they achieved better results. These methods consist of GA, GP and PSO where the most widely used method is GP algorithm because of its flexible tree-based representation that has the ability to deal with function operators in an easy way. However, although there are some works on EC methods for FC, various challenges continue to persist and their capability still needs to be explored completely. Researchers should also pay much attention to the hybridization of ECs and other methods which can help propose effective approaches for different types of problems.

Acknowledgements This work is supported by University Kebangsaan Malaysia, under grant number FRGS/1/2016/ICT02/UKM/01/2.

References

1. Russell, S., Norvig, P.: Artificial Intelligence: A Modern Approach, 3rd edn. Prentice Hall Press, USA (2009)
2. Blum, A.L., Langley, P.: Selection of relevant features and examples in machine learning. Artif. Intell. **97**(1–2), 245–271 (1997)
3. John, G.H., Kohavi, R., Pfleger, K.: Irrelevant features and subset selection problem. In: Proceedings of 11th International Conference on Machine Learning, pp. 121–129 (1994)
4. Quinlan, J.R.: C4.5: Programs for Machine Learning. Morgan Kaufmann Publisher (1993)
5. Liu, H., Motada, H.: Feature Extraction, Construction and Selection: A Data Mining Perspective. Kluwer Academic Publishers, Norwell (1998)
6. Xue, B., Zhang, M.: Evolutionary computation for feature manipulation: key challenges and future directions. In: Evolutionary Computation (CEC), pp. 3061–3067 (2016)
7. Xue, B., Zhang, M., Browne, W.N., Yao, X.: A survey on evolutionary computation approaches to feature selection. IEEE Trans. Evol. Comput. **20**(4), 606–626 (2016)
8. Liu, H., Yu, L.: Toward integrating feature selection algorithms for classification and clustering. IEEE Trans. Knowl. Data Eng. **17**(4), 491–502 (2005)
9. Tran, B., Zhang, M., Xue, B.: Multiple feature construction in classification on high-dimensional data using GP. In: SSCI, pp. 1–8 (2016)

10. Dai, Y., Xue, B., Zhang, M.: New representations in PSO for feature construction in classification. In: European Conference on the Applications of Evolutionary Computation, pp. 476–488. Springer, Berlin (2014)
11. Nguyen, H.B., Xue, B., Andreae, P.: A hybrid GA-GP method for feature reduction in classification. In: Asia-Pacific Conference on Simulated Evolution and Learning, pp. 591–604. Springer, Cham (2017)
12. Fogel, D.B.: Introduction to Evolutionary Computation. Wiley, New York (2007)
13. Vafaie, H., De Jong, K.: Genetic algorithms as a tool for restructuring feature space representations. In: Proceedings of 7th International Conference of Tools with AI, pp. 8–11 (1995)
14. Ahmed, S., Zhang, M., Peng, L., Xue, B.: Multiple feature construction for effective biomarker identification and classification using genetic programming. In: Proceedings of the 2014 Annual Conference on Genetic and Evolutionary Computation, pp. 249–256 (2014a)
15. Eiben, A.E., Smith, J.: From evolutionary computation to the evolution of things. Nature **521**(7553), 476 (2015)
16. Sondhi, P.: Feature construction methods: a survey. sifaka. uiuc. edu. **69**, 70–71 (2009)
17. Lillywhite, K., Lee, D.J., Tippetts, B., Archibald, J.: A feature construction method for general object recognition. Pattern Recogn. **46**(12), 3300–3314 (2013)
18. Drozdz, K., Kwasnicka, H.: Feature set reduction by evolutionary selection and construction. In: KES International Symposium on Agent and Multi-agent Systems: Technologies and Applications (pp. 140–149). Springer, Berlin (2010)
19. Shafti, L.S., Pérez, E.: Evolutionary multi-feature construction for data reduction: a case study. Appl. Soft Comput. **9**(4), 1296–1303 (2009)
20. Larsen, O., Freitas, A.A., Nievola, J.C.: Constructing X-of-N attributes with a genetic algorithm. In: GECCO Late Breaking Papers, pp. 316–322 (2002)
21. García, D., González, A., Pérez, R.: A two-step approach of feature construction for a genetic learning algorithm. In: Fuzzy Systems (FUZZ), pp. 1255–1262 (2011)
22. García, D., González, A., Pérez, R.: A feature construction approach for genetic iterative rule learning algorithm. J. Comput. Syst. Sci. **80**(1), 101–117 (2014)
23. Alfred, R.: DARA: data summarisation with feature construction. In: Second Asia International Conference on Modeling & Simulation, pp. 830–835. AICMS 08 (2008)
24. Firpi, H., Goodman, E., Echauz, J.: On prediction of epileptic seizures by computing multiple genetic programming artificial features. In: European Conference on Genetic Programming, pp. 321–330. Springer, Berlin (2005)
25. Chen, Q., Zhang, M., Xue, B.: Genetic programming with embedded feature construction for high-dimensional symbolic regression. In: Intelligent and Evolutionary Systems: The 20th Asia Pacific Symposium, IES 2016, Canberra, Australia, Nov 2016, pp. 87–102. Proceedings (2017)
26. Krawiec, K.: Genetic programming-based construction of features for machine learning and knowledge discovery tasks. Genet. Program Evol. Mach. **3**(4), 329–343 (2002)
27. Yazdani, S., Shanbehzadeh, J., Hadavandi, E.: MBCGP-FE: a modified balanced cartesian genetic programming feature extractor. Knowl.-Based Syst. **135**, 89–98 (2017)
28. Ahmed, S., Zhang, M., Peng, L.: A new GP-based wrapper feature construction approach to classification and biomarker identification. In: 2014 IEEE Congress on Evolutionary Computation (CEC), pp. 2756–2763 (2014b)
29. Elola, A., Del Ser, J., Bilbao, M.N., Perfecto, C., Alexandre, E., Salcedo-Sanz, S.: Hybridizing cartesian genetic programming and harmony search for adaptive feature construction in supervised learning problems. Appl. Soft Comput. **52**, 760–770 (2017)
30. Suganuma, M., Tsuchiya, D., Shirakawa, S., Nagao, T.: Hierarchical feature construction for image classification using genetic programming. In: IEEE International Conference on Systems, Man, and Cybernetics (SMC), pp. 001423–001428 (2016)
31. Tariq, H., Eldridge, E., Welch, I.: An efficient approach for feature construction of high-dimensional microarray data by random projections. PLoS ONE **13**(4), e0196385 (2018)
32. Tran, C.T., Zhang, M., Andreae, P., Xue, B.: Genetic programming based feature construction for classification with incomplete data. In: Proceedings of the Genetic and Evolutionary Computation Conference, pp. 1033–1040 (2017)

33. Muharram, M., Smith, G.: Evolutionary constructive induction. IEEE Trans. Knowl. Data Eng. **17**, 1518–1528 (2005)
34. Guo, H., Nandi, A.K.: Breast cancer diagnosis using genetic programming generated feature. Pattern Recogn. **39**(5), 980–987 (2006)
35. Liu, Q., Qiao, M., Sung, A.H.: Distance metric learning and support vector machines for classification of mass spectrometry proteomics data. Int. J. Knowl. Eng. Soft Data Paradig. **1**(3), 216–226 (2009)
36. Muharram, M.A., Smith, G.D.: Evolutionary feature construction using information gain and gini index. In: European Conference on Genetic Programming, pp. 379–388. Springer, Berlin (2004)
37. Neshatian, K., Zhang, M., Andreae, P.: A filter approach to multiple feature construction for symbolic learning classifiers using genetic programming. IEEE Trans. Evol. Comput. **16**(5), 645–661 (2012)
38. Liang, Y., Zhang, M., Browne, W.N.: Feature construction using genetic programming for figure-ground image segmentation. In: Intelligent and Evolutionary Systems, pp. 237–250. Springer, Cham (2017)
39. Tran, B., Xue, B., Zhang, M.: Genetic programming for feature construction and selection in classification on high-dimensional data. Memet. Comput. **8**(1), 3–15 (2015)
40. Ahmed, S., Zhang, M., Peng, L., Xue, B.: A Multi-objective genetic programming biomarker detection approach in mass spectrometry data. In: European Conference on the Applications of Evolutionary Computation, pp. 106–122. Springer, Cham (2016)
41. Cano, A., Ventura, S., Cios, K.J.: Multi-objective genetic programming for feature extraction and data visualization. Soft Comput. **21**(8), 2069–2089 (2017)
42. Xue, B., Zhang, M., Dai, Y., Browne, W.N.: PSO for feature construction and binary classification. In: Proceeding of the Fifteenth Annual Conference on Genetic and Evolutionary Computation Conference, GECCO, pp. 137–144 (2013)
43. Mahanipour, A., Nezamabadi-pour, H.: Improved PSO-based feature construction algorithm using feature selection methods. In: Swarm Intelligence and Evolutionary Computation (CSIEC), pp. 1–5 (2017)
44. Tran, B., Xue, B., Zhang, M.: Using feature clustering for GP-based feature construction on high-dimensional data. In: European Conference on Genetic Programming, pp. 210–226. Springer, Cham (2017)

Semantic Web Languages: Expressivity of SWL

Martin Žáček, Alena Lukasová, Marek Vajgl and Zdeňka Telnarová

Abstract The paper tries to discuss from a slightly higher level a focus oriented towards general properties that a Semantic Web Language (SWL) must have. We state as a ground the following two points of view: to compare (1) Expressivity of the SWL, and (2) a possibility to infer new knowledge from an SWL knowledge base with corresponding properties of the classical First Order Logics (FOPL) as a measure of their basic quality (now prepared). From expressivity, the language for the semantic web must be a common communication tool for computers as well as for people fulfilling an easy-to-use condition by means of adding more semantics directly into the language's syntax. After a discussion of properties and critical recommendations two languages, OWL DL1 and RDF CFL have been proposed here to become the SWL both having the expressivity comparable with the FOPL.

Keywords Semantic web language · SWL · First order logics · FOPL · Syntax · Semantic

1 Introduction

In our paper, we try to discuss from a slightly higher level with a focus oriented towards general properties that the Semantic Web Language (SWL) [1] or the Ontology Language (OL) of equivalent properties must have. On the base of evaluating a

M. Žáček (✉) · A. Lukasová · M. Vajgl · Z. Telnarová
Department of Informatics and Computers, Faculty of Science, University of Ostrava, 30. dubna 22, Ostrava, Czech Republic
e-mail: Martin.Zacek@osu.cz

A. Lukasová
e-mail: Alena.Lukasova@osu.cz

M. Vajgl
e-mail: Marek.Vajgl@osu.cz

Z. Telnarová
e-mail: Zdenka.Telnarova@osu.cz

M. Huk et al. (eds.), *Intelligent Information and Database Systems: Recent Developments*, Studies in Computational Intelligence 830,
https://doi.org/10.1007/978-3-030-14132-5_10

group of semantic web representation and communication means like formal ontology, description logics, semantic networks, RDF model and language, OWL language and RDF CFL language we state as basics the following two points if view.

First, it seems necessary to compare an expressivity and users' usability of the SWL languages in discussion with the corresponding properties of the classical First Order Logics (FOPL) [2] as a measure of the basic quality of formal languages.

Second, the language must use some of the methods FOPL of formal reasoning to be able to inferring further knowledge from a knowledge base [3]. The language must be the common communication tool for computers and also for people by fulfilling the easy-to-use condition—adding more semantics direct into the language syntax.

A further argument that speaks to the benefit of using one of the best results of a long-time development of the FOPL is the fact that there exist series of new algorithms for solving concrete tasks. Some of them ought to be usable also within SWL languages.

The paper [4] follows in an excellent stream of facts and their mutual connections some historical roots of the language OWL that became until now the main communication tool for a generally Semantic Web Language (SWL). Our article tries to make a possibility of choosing a bit richer. Into the genealogical tree leading towards OWL we have pasted a new branch about the RDF model and language completed by quantifiers that forms together with Richards CFL an alternative language RDF CFL possible to fulfill all basic demands to become a real SWL.

Even if the development of a semantic web language was until now much more complicated with a taking part of a lot of people or companies the authors of the article try to contribute to an analyzing of the actual state by drawing it as a cloud of demands, streams or conditions and events that at the end would have brought into the world easy-to-usable semantic web languages having all the demands fulfilled.

2 An Idea of Semantic Web

Natural language is adequate for communication/disputation between people, but not suitable to do it by computers. The idea of the semantic web adds to the human recognizing of the world a new demand: besides human readers and users it must recognize and use knowledge and information in the web also by computers. The demand also lays a new condition on a Semantic Web Language (SWL): to be readable and recognizable to people as well as to computers.

The name "Ontology Language" (OL) [5] or "Ontology Web Language" (OWL) [6] appeared in literature before the name "Semantic Web Language" but both ought to fulfill a lot of common demands. Because of the complexity of the description, all those languages need to have a special domain of knowledge. SWL must fulfill one demand more "to be easy-to understandable for people"—it means to have got some semantic elements included directly within the language syntax.

During the last decades, it has been proposed several languages, some of them based on FOPL or XML. It is a question which properties ought to have the main common features of the SWL. In order to translate a statement into such a suitable form minimal the following conditions are to be fulfilled:

1. A necessity to introduce a system of unique identifications for the statement components within a domain of interest in a domain ontology, by which they could become elements of an automated process. Artificial intelligence (AI) sees the domain ontology as a framework fulfilling the role of a working model for entities, their hierarchical ordering and their interactions in some of the particular domain of discourse. The way to general principles of building domain ontologies we can see today as a composition of various directions each of them preferred its own goals or preferences. Nevertheless, they have formed together with a unique stream with many interesting branches like it is shown in a graph (semantic network) at Fig. 1.
2. To have got the property of clear readability and usability of statements by people as well as by computers. To fulfill this condition, it is necessary to complete statements by their meanings, all in a corresponding explicit formal way.
3. It has become clear from the user point of view that some simplification of the basic knowledge about the domain to be represented had to take place. It had been there a necessity to begin at the conceptual level of knowledge representation with a new paradigm of an abstract seeing a domain to be described.
4. From the very beginning of web technology, there has been a special demand for the web quality named "open world assumption" followed the fact of continual development of a knowledge base collected on the web.

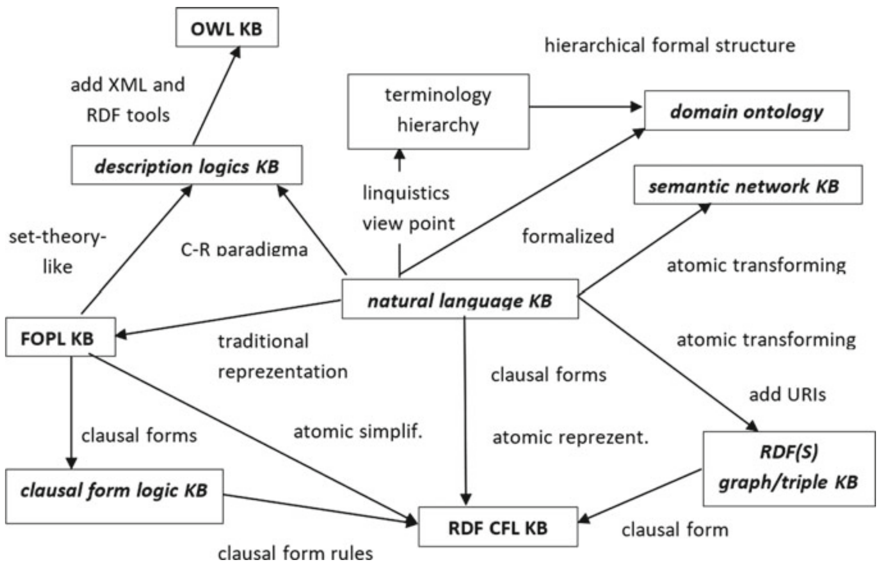


Fig. 1 Semantic network of a ground part of article's vocabulary

3 From Natural Language and FOPL Towards a Real Semantic Web

3.1 Expressive Power Like at the FOPL

Propositional and later especially the predicate logics accompanying development of theoretical mathematics has brought one for artificial intelligent (AI) ground tool in its capability to represent knowledge formally with a guarantee of high expressivity near that one of a natural language. A full description of all the important facts within a domain in question in the language of First Order Predicate Logics (FOPL) have been taken by AI specialists as a best usable correct picture with a maximal expressivity and factual reliability of representation. Standard syntax and the model-theoretic (MT) [7] semantics of the FOPL language or both together have served to people user a complete information. For transfer the information to computers there have been XML-like tools. But computers users could have got a meaning of information only with a help of IT specialists.

Nevertheless, the full traditional knowledge representation in the mathematic-logics based FOPL language has brought a contribution to the goal by the high quality of expressivity and possibilities of precise descriptions. Later new languages could have been compared with the FOPL as a guarantee of high quality.

3.2 A Common Base—The Concept-Relationship Paradigm of Seeing a World

At the beginning, a formal modeling process a developer uses to start by a conceptualization. It means a special seeing a structure of a domain of his interest using a special paradigm. Concept-Relationship (C-R) paradigm seems to be a tool of how to make the representation simpler with sharing all the original factual content.

Within the (C-R) paradigm the world consists of concepts that have their properties and take place in various connections or relationships with another concept. Concepts, their properties, and relationships serve a general ground for a possibility to express a domain of interest in the form of the domain ontology.

The required property of expressivity (point 1 above) relative to the corresponding FOPL and at the same time the syntax simplification (point 3 above) has been achieved using a small decomposition of participated predicates to unary predicates for representation of concepts and binary predicates for representation of their relationships. Then it was easy to customize a binary predicate to represent an atomic statement in the form of graph vectors, in which the edge representing the predicate connects its two nodes labeled by the attributes of the binary predicate. On this base has been created the whole series of domain-oriented ontologies, i.e., terminological dictionaries of many domains, specifying concepts and their relationships in their contexts and the relationships to hierarchical structures.

In a bit more general point of view, the RDF graphics [8], as well as semantic network of a concept represented by graph, give a possibility to see a concept together with an explanation of their semantics. The concept shown along with accompanying information about its context of a modeling world has been defined in this way by its "intentionally defined" semantics. The name "associative network" is also appropriate as these networks support understanding of the importance of the concept about the concepts that form its importance.

"Domain Ontology" [9] is just another name for the special vocabulary of the domain in question. In the context of the Semantic Web, ontologies are expected to provide structured vocabularies that explicate the relationships between different terms, allowing intelligent agents (and also humans) to interpret their meaning.

The graph at Fig. 1 explains using graphic representation in a semantic network a ground part of a vocabulary that takes part in our article. It shows a knowledge base (KB) in a natural language, its formal representation in the FOPL, its modification CFL (Clausal Form Logic) [10] and two of the semantic web languages OWL and RDF CFL [11].

For example, "traditional representation" makes from a "natural language KB" a "FOPL KB" or "atomic transforming" rewrites a "natural language KB" into a "semantic network KB" or the "set-theory-like" formalizing makes from the "FOPL KB" the "DL KB"

Next two paragraphs specify principles of representation in the case of the OWL and RDF CFL [12] from the FOPL.

4 Proposal of RDF CFL as an SWL

4.1 RDF Model as a Main Principle of a Semantic Web Language

A main idea of the RDF model was at the beginning similar as that one of the approaches of semantic networks—seeing and writing knowledge statements about the world just by rewriting the arguments from natural language into a set of triples or graphic easy-to-understand form corresponding to the set of elementary formulas of the FOPL. Taking the world of discourse under the general concept-relations paradigm led the formal representation towards a crucial simplified form of elementary statement about the world. Thus, the language naturally followed the RDF(S) [8] model.

For example, at a simple sentence "Each man that has been married to a wife is her husband." In the natural language there exists more than a unique traditional representation method of rewriting the sentence above into the FOPL form like for example

$$\forall x(\text{man}(x) \,\&\, \exists y(\text{woman}(y) \,\&\, \text{event}(x, y, \text{wedding})) \rightarrow \text{husband}(x, y))$$

The required property of expressivity relative to the corresponding FOPL and at the same time the syntax simplification has been achieved using a small decomposition of a participated predicate event(x, y, wedding) to a binary predicate married(x, y) of the corresponding meaning. Now the expression **(E FOPL)** uses only unary predicates for concepts and binary predicates for relationships.

$$\forall x(\text{man}(x) \,\&\, \exists y(\text{wife}(y) \,\&\, \text{married}(x, y)) \rightarrow \text{husband}(x, y))(\textbf{E FOPL})$$

4.2 RDF CFL as an SWL Language

Development in the formal knowledge representation clearly shows that the language of the FOPL and its *Clausal Form Logic* (CFL) [10] specifically is the appropriate formal language that can virtually represent any assertion formulated in a natural language Moreover, the "if-then" statements (conditional statement—*clauses*) in the CFL say that a *consequent* composed as a disjunction of some predicate atoms follows from *antecedent* composed as a conjunction of some predicate atoms.

Atomic statements $P_1, \ldots, P_m, Q_1, \ldots, Q_n$ generally form the following structure of the clause:

$$P_1 \,\&\, \ldots \,\&\, P_m \rightarrow Q_1 \vee \ldots \vee Q_n \text{ or}$$
$$P_1, \ldots, P_m \rightarrow Q_1, \ldots, Q_n \text{ or simply } \langle\text{antecedent}\rangle \rightarrow \langle\text{consequent}\rangle,$$

where $P_1, \ldots, P_m, Q_1, \ldots, Q_n$ are positive binary predicate atoms (vectors).

In the RDF CFL, a special representation of a knowledge atom is used: an RDF triple or a vector in the case of its corresponding graph language, which is already very well established in associative networks.

Our aim is to propose here a formal language RDF CFL (defined for example at [12]) for a representation and manipulation of knowledge gathered on the web site as one of the SWL languages, as it

1. has a high expressivity relative to natural languages, similarly to the case of the FOPL,
2. has a property, knowledge represented in the language being placed into some external contexts, in order to capture better its semantics,
3. uses a relatively simple language syntax, allowing a machine readability and handling.
4. The formal language with the characteristics above naturally ought to provide a mechanism of automated deduction in the frame of a semantically correct and complete formal method of resolution.
   When the idea of the Semantic Web has appeared at the beginning of this Millennium, the file of the properties shaped above has got one requirement more:

5. The language of knowledge representation should be a web language that ought to be suitable to maintain knowledge by people as well as by machines.

5 OWL Ontology Language as an SWL

A major influence on the design of OWL was the requirement to maintain the maximum upwards compatibility with existing web languages (see [3]). Description logic has brought the expressive and reasoning power to the OWL language of the semantic web. Using RDF model extended by the RDFS the semantic web has become an ontology language that can formally describe the meaning of terminology used in web documents. OWL now contains a broad set of elements through which it is possible to create ontologies from the simplest to the most sophisticated structures built on unlimited use of RDF constructs.

The standard technique for specifying the meaning of the description logics ought to be via a model-theoretic semantics, whose purpose is to explicate the relationship between the language syntax and the intended model within a special domain (according to a part of the definition).

5.1 Description Logic Languages DL and DL1

Authors of the [3] believe that a key feature of DL is that one based on the syntax and well-defined semantics expressed. However, a serious mistake was found behind some experimental results of applying the DL semantic tableau reasoner. It has been shown that the ground of the mistakes consists in a using of the universally quantified constructor $\forall R \cdot C$ with its meaning defined as $\{\mathrm{x} \in \Delta^{I} | \forall\, \mathrm{y}(R(\mathrm{x}, \mathrm{y}) \to C(\mathrm{y}))\}$.

If the concept $\forall R \cdot C$ ought to do a further restriction (specification) of that one defined by $\exists R \cdot C$ the extents of them must always (in all interpretations) fulfill a subset inclusion of extent of the concept $\exists R \cdot C$. It implies the following axiom must hold:

$$\forall R \cdot C \sqsubseteq \exists R \cdot C \tag{1}$$

The mistakes of original DL [1] becomes from a using of universally quantified constructor $\forall R \cdot C$ with its meaning defined as $\{\mathrm{x} \in \Delta^{I} | \forall \mathrm{y}(R(\mathrm{x}, \mathrm{y}) \to C(\mathrm{y}))\}$. That leads to a conclusion that the formula (1) $\forall R \cdot C \sqsubseteq \exists R \cdot C$ does not hold. But the concept expressions $\forall R \cdot C$ and $\exists R \cdot C$ belong to the tools for possibilities of expressing a hierarchy of concept extensions. The reason lays in their set-theoretic meaning:

If the concept $\forall R \cdot C$ ought to do a further restriction (specification) of that one defined by $\exists R \cdot C$ the extents of them must always (in all interpretations) fulfill a subset inclusion of extent of the concept $\exists R \cdot C$.

The following analysis shows why it is necessary to replace the language DL by the language DL1 with another semantics of the $\forall R \cdot C$ constructor.

5.2 Syntax and Semantics of DL and DL1 Languages

Let us have a set of concept symbols (unary predicate symbols), a set of role symbols (binary predicate symbols), a set of individual object name symbols and a set of variable names. Formulas of a language DL/DL1 are concept formulas (Cformulas) or role formulas (Rformulas), both created on the sets of symbols above by a special set of DL/DL1 constructors.

Let us analyze extents of quantified constructors $\exists R \cdot C$ and $\forall R \cdot C$ to see their expressive power and their capability to express concept subsumption (1).

The extent of the concept $\exists R \cdot C$ with the semantics defined in the book [1] as

$$(\exists R \cdot C)^I = \{\mathrm{x} \in \Delta^I \mid \exists \mathrm{y}\ (R(\mathrm{x}, \mathrm{y})\ \&\ C(\mathrm{y}))\} \tag{2}$$

must be a restriction of the extent of the concept $\exists R \cdot \mathrm{T}$. The extent of $\exists R \cdot C$. involves all the objects x of the universe of discourse for which an object o (o $\in$ ext(C)) exists within the extent of the role R such that R(x, o) holds.

The extent of the concept $\forall R \cdot C$ must always be a further restriction of the extent of the concept $\exists R \cdot C$.

However, the extent of the concept $\forall R \cdot C$ within the semantics is defined in [1] as

$$(\forall R \cdot C)^I = \{\mathrm{x} \in \Delta^I \mid \forall \mathrm{y}\ (R(\mathrm{x}, \mathrm{y}) \rightarrow C(\mathrm{y}))\} \tag{3}$$

also involves all those objects x of the universe of discourse that fulfills $\neg R$ (x, y).

As we do not want to use in our practice such a constructor $\forall$R $\cdot$ C with the meaning contrasting with a commonsense understanding we have introduced a description logic DL1 with such a semantics that make the extent of the concept $\forall$R $\cdot$ C to be a further restriction of the extent of the concept $\exists$R $\cdot$ C, so the extent of the concept $\forall$R $\cdot$ C is a subset of the extent of the concept $\exists$R $\cdot$ C. It implies the axiom (1) must hold.

The logical validity of **(1)** is easy-to proof with the help of the corresponding FOPL transcription that specifies an implication of the meanings of both quantified constructors of description logics.

The following subsumption corresponding to **(1)** holds in the FOPL language and has a form of implication

$$\forall x(\forall y(R(x, y)\,\&\,C(y)) \rightarrow \exists y(R(x, y)\,\&\,C(y)))\ (1\ \text{FOPL})$$

It is easy to prove that (**1** FOPL) holds by a semantic tableau tree of the negation (**1** FOPL) with a symbol of coach (×) at the end of each of branches.

Now let us reason two semantic tableaus of the implication **(1′)** and (1″) in the corresponding transcriptions of the **(1)** of the DL and the DL1 according to expressions characterizing both of their semantics of $\forall R \cdot C$.

Indirect proof of the validity of the **(1′)** transcription of the DL axiom **(1)** (Fig. 2):

The semantic tableau of the **¬(1′)** can't lead to false. So the implication (1′) at the headline of the proof cannot be a true logical implication of the FOPL.

The last of the semantic tableau proofs shows in the FOPL transcription a semantic tableau using definition of the semantics of $\forall R \cdot C$ in the DL1 language (Fig. 3):

¬∀x (∀y (R(x, y) → C(y)) → ∃y (R(x, y) & C(y))) **(1′ FOPL)**

∃x ¬ (∀y (R(x, y) → C(y)) → ∃y (R(x, y) & C(y)))

¬ (∀y (R(a, y) → C(y)) → ∃y (R(a, y) & C(y)))

∀y (R(a, y) → C(y)), ¬∃y (R(a, y) & C(y))

∀y (R(a, y) → C(y)), ∀y ¬ (R(a, y) & C(y))

R(a, b) → C(b), ¬R(a, b) ∨ ¬C(b)

¬R(a, b) ∨ ¬C(b), ¬R(a, b) ¬R(a, b) ∨ ¬C(b), C(b),

¬R(a, b) ∨ ¬C(b), C(b) ¬R(a, b) ∨ ¬C(b)

¬R(a, b), C(b) ¬C(b),C(b) ¬R(a, b) ∨ ¬C(b)

¬R(a, b), C(b) ¬C(b), C(b)

Fig. 2 Semantic tableau of the (**1′**) expression

∀x ¬ (∃t R(x,t) &∀y (R(x, y) → C(y)) → ∃y (R(x, y) & C(y))))

(1′′) with the DL1semantics

∃x ¬ (∃t R(x,t) & ∀y R(x, y) → C(y))) → ∃y (R(x, y) & C(y)))

¬ (∃t (R(a,t) & ∀y (R(a, y) → C(y))) → ∃y (R(a, y) & C(y)))

R(a,k) & ∀y (R(a, y) → C(y)) → ∀y ¬(R(a, y) & C(y)))

(R(a,k) & (R(a, k) → C(k))), ¬ (R(a, k) & C(k))

R(a,k), R(a, k) → C(b), ¬R(a, b) ∨ ¬C(b)

¬R(a, k) ∨ ¬C(k), R(a,k), ¬R(a, k) ¬R(a, k) ∨ ¬C(k), R(a,k), C(k)

¬R(a, k), R(a,k), C(k) ¬C(k),R(a,k), C(k)

Fig. 3 Semantic tableau of the (**1″**) expression

5.3 An Alternative Proposal and Negation of the ∀R · C Semantics in DL1

Authors of the [3] introduced rewriting rules for negations of the concepts $\exists R \cdot C$ and $\forall R \cdot C$ within the language DL. It means the following equivalencies hold in the book [1]:

$$\neg\exists \mathrm{R} \cdot \mathrm{T} \Leftrightarrow \forall\neg \mathrm{R} \cdot \mathrm{T} \tag{2}$$

$$\neg\exists \mathrm{R} \cdot \mathrm{C} \ \Leftrightarrow \forall \mathrm{R} \cdot \neg \mathrm{C} \tag{3}$$

$$\neg\forall \mathrm{R} \cdot \mathrm{C} \Leftrightarrow \exists \mathrm{R} \cdot \neg \mathrm{C} \tag{4}$$

However, the equivalencies (2)–(4) hold only in the world where the basic relation (1) between quantified constructors ∃R · C and ∀R · C does not hold.

We infer the corresponding rewriting expression (4′) of the negation on the base of the semantics of the ∀R · C:

$$\begin{aligned}(\forall R \cdot C)^I &= \{\mathrm{x} \in \Delta^I | \neg(\exists \mathrm{t} R(\mathrm{x}, \mathrm{t}) \,\&\, \forall \mathrm{y}(R(\mathrm{x}, \mathrm{y}) \rightarrow C(\mathrm{y})))\} \\ &= \{\mathrm{x} \in \Delta^I | \forall \mathrm{t} \neg(R(\mathrm{x}, \mathrm{t}) \vee \exists \mathrm{y} \neg(R(\mathrm{x}, \mathrm{y}) \rightarrow C(\mathrm{y}))\} \\ &= \{\mathrm{x} \in \Delta^I | \forall \mathrm{t} \neg R(\mathrm{x}, \mathrm{t}) \vee \exists \mathrm{y}(R(\mathrm{x}, \mathrm{y}) \,\&\neg C(\mathrm{y}))\} \\ &= (\forall\neg R \cdot \mathrm{T} \sqcup \exists R \cdot \neg C)^I\end{aligned}$$

$$\neg\forall R \cdot C \Leftrightarrow \exists R \cdot \neg C \ \sqcup \ \forall\neg R \cdot \mathrm{T} \tag{4′}$$

Similarly, we infer the corresponding rewriting expression (3′) of the negation on the base of the semantics of the $\exists R \cdot C$:

$$\begin{aligned}(\neg\exists R \cdot C)^I &= \{\mathrm{x} \in \Delta^I | \forall \mathrm{y} \neg(R(\mathrm{x}, \mathrm{y}) \,\&\, C(y))\} \\ &= \{\mathrm{x} \in \Delta^I | \forall \mathrm{y} \neg(R(\mathrm{x}, \mathrm{y}) \,\&\, (R(\mathrm{x}, \mathrm{y}) \rightarrow C(y))\} \\ &= \{\mathrm{x} \in \Delta^I | \forall \mathrm{y}(\neg R(\mathrm{x}, \mathrm{y}) \vee \neg(R(\mathrm{x}, \mathrm{y}) \rightarrow C(\mathrm{y}))\} \\ &= \{\mathrm{x} \in \Delta^I | \forall \mathrm{y}(\neg R(\mathrm{x}, \mathrm{y}) \vee (R(\mathrm{x}, \mathrm{y}) \,\&\, \neg C(\mathrm{y}))\} \\ &= (\forall\neg R \cdot \mathrm{T} \sqcup \forall R \cdot \neg C)^I\end{aligned}$$

$$\neg\exists R \cdot C \Leftrightarrow \forall\neg R \cdot \mathrm{T} \sqcup \forall R \cdot \neg C \tag{3′}$$

As the R(x, y) in the second disjunction implies ∃t R(x, t) and ¬R(x, y) in the first disjunction must be true the following implication must hold (Table 1):

$$\neg\exists \mathrm{R} \cdot \mathrm{C} \Leftrightarrow \forall\neg \mathrm{R} \cdot \mathrm{T} \sqcup \forall \mathrm{R} \cdot \neg \mathrm{C}$$

Table 1 Rewriting rules for negative quantified formulas in DL1

Negative quantified formula	Rewriting rule
$\neg\exists R \cdot \top$	$\forall\neg R \cdot \top$
$\neg\exists R \cdot C$	$\forall\neg R \cdot \top \sqcup \forall R \cdot \neg C$
$\neg\forall R \cdot \top$	$\exists\neg R \cdot \top$
$\neg\forall R \cdot C$	$\forall\neg R \cdot \top \sqcup \exists R \cdot \neg C$

5.4 The Language OWL DL1 Shares Properties of the FOPL

In the case of description logics DL [3], the C-R paradigm probably at the first time has taken its leading place. Until now there has been only a natural demand to have a simplified language with sharing a high expressivity like in the case of the FOPL. Both demands have been fulfilled as a simplified version of the FOPL at both description logics DL or DL1 on the base of the concept-relationship construction. At both of the cases, the simplification consists in the replying formal quantified specifying of sets by their inclusion relationships. Here a pure formal logics approach becomes more or less a set-theory approach that makes readability and understandability of formalisms a bit less difficult.

Expressivity of a language of ontology as well as of the SWL should be comparable with the language of the FOPL with a well formally specified semantics. Ontology building used to go along tests of consistence of the growing set of concepts or relationships: concept hierarchy as well as their synonyms ought to be taken into account a possibility of integration of ontologies can be supported by inter-ontology relationships and testing for consistency and computing the integrated concept hierarchy.

In description logic, an ontology developer defines important notions of domain ontologies, especially classes or relations of them regarding concepts and roles. The concepts of the DL/DL1 language are represented by unary predicates and roles by binary predicates. Further constructors in a quantified form give in the DL/DL1 possibilities to restrict concepts or roles to their various specifications. It has appeared there that demand for building a new language on the base of the DL/DL1 atomic elements has also been recognized by human minds.

The authors of the basic DL [3] have been sure that the ontology language DL supports maximum expressiveness while retaining computational completeness and decidability. We discussed the problem in our article [13] and had a proposal to make in our DL1 version some correction on the definition of the DL syntax.

The OWL DL1 with the description logic DL1 of our definition has the properties of soundness and completeness from the semantic point of view and belongs to the semi-decidable languages corresponding to the FOPL.

It seems very rational to build OWL on the base of DL1, not on the DL. The question of decidability would be a bit less complicated; special lighter conditions do not have to be defined. For example, DL1 does not assume the "unique name assumption." Three level of OWL (OWL-lite, OWL-DL, and OWL-full become useless, OWL-DL1 corresponding with the FOPL must be undecidable. However, there exists at least one DL1 procedure (see further paragraphs) that guaranties a semi-decidability of the OWL-DL1.

5.5 Comparing Properties of OWL DL1 and RDF CFL as SWL

From the FOPL, the language simplification in both SWL languages OWL and RDF CFL have in general the same approach using the C-R paradigm but a distinguishing feature of their simplifying.

The language OWL with its main part DL simplifies the FOPL simply by an own set-theory principle of the C-R paradigm in the conceptualized world of concepts and relationships.

The language RDF CFL has to simplify the FOPL also using a semantic partition of knowledge into its corresponding set of knowledge atoms within the world of C-R concepts and relationships. Both of them are only instances of a special formalization.

5.6 The Composition of Semantics Direct into Syntactic Forms of Representation

The associative (semantic) network that has brought some specialists in English linguistic was a first attempt to compose semantics into syntactic forms of representation. Syntactic networks have been used there for support of (people, not computer) a better understanding of information content. The necessity of graphic drawing knowledge has brought a new direction of portioning information into bags of linked atomic statements according to real connectivity of their meanings. However, it has also come a time to solve a further problem—adding a correct meaning to a concrete knowledge direct into the language of formal representation. The correct meaning to the formal elements of knowledge has brought the model and formal language of RDF using unique pointing towards explicit specifications of concrete meanings that have been written in a domain ontology for each one of the syntax elements of the knowledge.

6 Conclusion

As the aim of the first part of the treatise about web-language candidates for a role of an SWL has been chosen their capability to share of high expressivity of the FOPL after a necessary simplifying and to preserve the expressivity of the original FOPL language at the same time. The paper analyzes and compares tools of making the presented languages OWL DL1 and RDF CFL simpler in their syntax and ordering a bit more semantics direct into their new syntax The OWL DL1 reaches the goal by replying formally quantified specifying of sets by their inclusion relationships. So a pure formal logics approach becomes more or less a set-theory approach that makes readability and understandability of formalisms a bit less difficult at the case of the RDF CFL language simplifying means to decompound knowledge into badges of elementary statements expressing together the same meaning as the original knowledge.

The paper more or less builds on the paper [4] that has been drawn in an excellent stream of facts, their mutual connection principles and some historical roots the language OWL (Ontology Web Language) that have become until now the main communication tool for a general OL or SWL. We add to the tree of history and facts a stronger part of a very important branch—the RDF model completed by quantifiers that gives a possibility of using the RDF principle not only in the role of creator of a rich semantic variability but together with resolution principle in Clausal Form Logic (CFL) also a role of a language that has been able to fulfill all the demands to become a SWL.

The clausal form logic (CFL) has originated from the FOPL and belongs now to ground approaches of logical programming. Together with a precise semantics of the RDF model and the possibility of corresponding graph representation our formal language RDF CFL belongs to simple and easy-to-use SWL languages. Inference system based on the resolution rule shares properties of semantic correctness and completeness with the FOPL. Moreover, the formal system with the principle of the CFL and with its resolution rule of formal deduction has the system the property of semidecidability.

In the article, we focus ourselves on nowadays approach used for current common formalisms of deduction within the idea of semantic web from expressivity equivalent to FOPL knowledge representation.

Knowledge representation and realizing an inference has been discussed namely within that one based on the description logic DL1 and that one based on the RDF CFL modeling principles. We also present a discussion how some of the current approaches lack properties important for such modeling. Some possibilities of alternate ways have been suggested considering a nowadays popular form of linked data.

References

1. Straccia, U.: Foundations of Fuzzy Logic and Semantic Web Languages. Chapman and Hall/CRC (2016)
2. Smullyan, R.R.: First-Order Logic, vol. 43. Springer Science & Business Media (2012)
3. Baader, et al. (eds.): The Description Logic Handbook—Theory, Interpretation, and Applications. Cambridge University Press (2003)
4. Horrocks, I., Patel-Schneider, P.F., Van Hamelen, F.: From SHIQ and RDF to OWL. The making of a web ontology language. J. Web Semant. (2003). https://doi.org/10.1016/j.websem.2003.07.001

5. McGuinness, D.L., Van Harmelen, F.: OWL web ontology language overview. W3C Recomm. **10**(10) (2004)
6. Hastings, J.: Primer on ontologies. In: The Gene Ontology Handbook, pp. 3–13. Humana Press, New York (2017)
7. Barwise, J., Feferman, S., Feferman, S. (eds.): Model-Theoretic Logics, vol. 8. Cambridge University Press (2017)
8. Lukasová, A., Vajgl, M., Žáček, M.: Reasoning in RDF graphic formal system with quantifiers. In: Proceedings of the International Multiconference on Computer Science and Information Technology, Poland, pp. 67–72. IEEE Computer Society, Mragowo (2010)
9. Maedche, A., Staab, S.: Ontology learning for the semantic web. IEEE Intell. Syst. **16**(2), 72–79 (2001)
10. Richards, T.: Clausal Form Logic. An Introduction to the Logic of Computer Reasoning. Addison-Wesley (1989)
11. Lukasová, A., Žáček, M., Vajgl, M.: Reasoning in Graph-based clausal form logic. IJCSI Int. J. Comput. Sci. Issues **9**, 37–43 (2012). ISSN 1694-0814
12. Lukasová, A., Žáček, M., Vajgl, M., Kotyrba, M.: Resolution reasoning by RDF clausal form logic. IJCSI Int. J. Comput. Sci. Issues **9**(3) (2012). ISSN (Online): 1694-0814
13. Lukasová, A., Žáček, M., Vajgl, M.: Building a non-monotonic default theory in GCFL graph-version of RDF. In: Modern Trends and Techniques in Computer Science, AISC 285, pp. 455–466. Springer International Publishing, Springer International Publishing, Switzerland (2014)

Application of Distance Measurement NLP Methods for Address and Location Matching in Logistics

Leo Mrsic

Abstract Paper is based on research based on linguistic terms denoting an address used for delivery services in logistic. Distance measurement NLP methods are widely usable in text mining and can be used to find the similarity among sentence or document. As part of logistics process, being able to determine correct address using machine learning we need to tackle issue of two addresses comparison (street name, city name etc.) is crucial for efficient service. This paper explains comparison techniques based on similarity score that can be calculated using distance measurement. As part of process, several distance measurements were compared while conclusion include results and recommendation on usage in address and location matching in logistics (post services).

Keywords Distance measurement NLP methods · Address matching in logistics · Location matching in logistics · Text mining

1 Introduction

1.1 Business Perspective

Business optimization and constant work on utilizing technology advancement is one of the top priorities of successful business. Development and availability of technology with timely engagement around the overall refinement portfolios of services by using data, opens up a completely new dimension in which data becomes “the most valuable resource of the company” and by means of them they create tools that through new business models improve all business channels and significantly contribute to strategic goals. This paper introduces automatic address routing as part of express delivery process in logistics applicable for daily matching services or quality assurance processes [1–5] Automatic routing process include pairing the

L. Mrsic (✉)
Algebra University College, Ilica 242, Zagreb, Croatia
e-mail: leo.mrsic@algebra.hr

M. Huk et al. (eds.), *Intelligent Information and Database Systems: Recent Developments*, Studies in Computational Intelligence 830,
https://doi.org/10.1007/978-3-030-14132-5_11

address between sender and receiver from the transaction system, address database with machine learning elements and the option to systematically learn through a mismatched return address. Within the business processes of express delivery several systems can benefit from proposed solution: automatic routing of an espresso delivery, inadequate or missing address breaks, dynamic routing of express delivery, more adequate arrangement of fixed routes, optimization and predicting the time required for each group to deliver shipments, statistics by individual groups and phases (e.g. “at the time of delivery shipment for the exact day”) [6–8].

1.2 Research Environment

As part of this research proposed model was developed in environment of approximately 5,000,000 shipments which are handled annually in the express delivery system. For this particular project researchers had access to partial data taken from two-year history of matching statistics or unsorted shipments from the earlier period that was used for learning. Data in paired addresses would contain a degree of relationship-based security based on experience where those addresses that are not labeled as a “negative” system after some time puts them in properly paired [9–11] (Fig. 1).

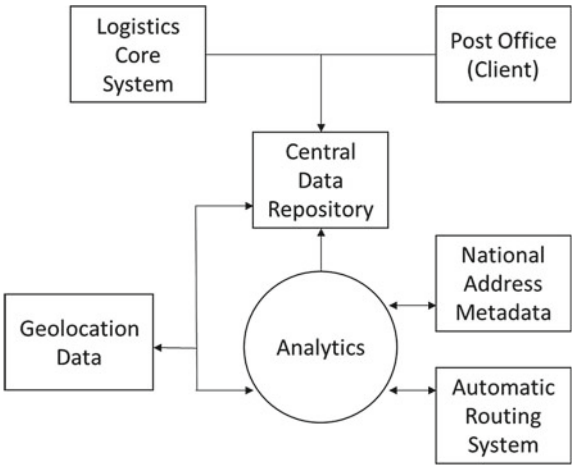


Fig. 1 Research environment

2 Application of Distance Measurement NLP Methods

2.1 String Distance Measurement

Different researchers from mathematics or computer science are familiar with "string metric" (also known as a function of string similarity metric or string distance). Term and technique is commonly used to measure distance between two text strings for approximate string matching or comparison in fuzzy string searching. As a prerequisite for a string metric is fulfillment of the, so called, triangle inequality. For example, the strings "Leo" and "Leonard" can be considered to be "close". As a result of string metric equitation, process generate a number indicating an algorithm-specific indicator/value of distance [12–16] (Table 1).

One of the most widely known string metrics is called the Levenshtein distance (familiar under "edit distance" name). Like others, it operates between two input strings, returning a value equivalent to the number of substitutions and deletions needed in order to transform one input string into another. Within its additions, string metrics expanded to include phonetic, token, grammatical and character-based methods of statistical comparisons. String metrics are used heavily in various industries and for various purposes including fraud detection, fingerprint analysis, plagiarism detection, ontology merging, DNA analysis, RNA analysis, image analysis, evidence-based machine learning, database data deduplication, data mining, incremental search, data integration, and semantic knowledge integration [17–20].

Table 1 String distance measurements comparison

Selected string distance measures	Example
Euclidean distance	Comparing the shortest distance among two objects while score means the distance between two objects (0 = identical)
Cosine distance	Determine the angle between two objects is the calculation method to the find similarity (1 means that they are same in orientation)
Jaccard similarity	The measurement is referred to number of common words over all words. Jaccard Similarity = (Intersection of A and B)/(Union of A and B) (1 means that they are identical)
Levenshtein distance	kitten and sitting have a distance of 3. kitten → sitten (substitution of "s" for "k") sitten → sittin (substitution of "i" for "e") sittin → sitting (insertion of "g" at the end)
Hamming distance	"karolin" and "kathrin" is 3
Most frequent k characters	MostFreqKeySimilarity ('research', 'seeking', 2) = 2

2.2 String Similarity Measurement Using Distance

Using multiplication from string, similarity among sentence or document is more attractive however more challenging we well. Text is different than a number therefore we cannot compare the difference between "Lemon" and "Mango" but similarity score has been calculated. There are several methods with assumption that it can deliver similarity however they do not handle synonym scenario neither synonyms [21, 22]. Linking addresses to/from the registries and address database is recognized as the problem of machine translation. Research include preparation and coaching of this part of the system which was tested and validated over existing data before production. During development, the use of recurrent neural networks of Long Short-Term Memory was used for the needs of specific content (addresses with all address details) will be recruited for translation of the address string with the help of fuzzy strings of characters of the central records in a string from the metadata repository. Once developed, network will result in a candidate/link to metadata pair in the first step for which the second step will skew the degree of coincidence due way of connecting. A pure fuzzy connection would be useful for each character string from base addresses searched for N matching string from metadata records, but here the problem is reversed, and close candidates from the metadata we are asked for the default string. With matching candidates' likelihood of matching will be displayed and the candidates will be sorted by probability matching. This procedure formed more most likely outputs of match related probabilities [23–25]. The system will be able to receive the data on the dispatches/pairs for which on-time delivery was not possible and used as a basis for knowledge for future pairing. Also, events where paired address does not occur was marked as a "negative" in system, while the links was be stored in the internal learning base as appropriate paired data sequences and learn from their content and links. Improving self-discipline precision involves adding "attention" mechanisms, residual bonds and, where appropriate, other deeper machine learning techniques translation but also supported in depth learning.

2.3 Levenshtein Distance

In information theory, linguistics and computer science, the Levenshtein distance is a string metric for measuring the difference between two sequences. Informally, the Levenshtein distance between two words is the minimum number of single-character edits (insertions, deletions or substitutions) required to change one word into the other. It is named after the Soviet mathematician Vladimir Levenshtein, who considered this distance in 1965.

Table 2 Levenshtein string distance measurement example (words "Saturday" and "Sunday")

		S	a	t	u	r	d	a	y
	0	1	2	3	4	5	6	7	8
S	1	0	1	2	3	4	5	6	7
u	2	1	1	2	2	3	4	5	6
n	3	2	2	2	3	3	4	5	6
d	4	3	3	3	3	4	3	4	5
a	5	4	3	4	4	4	4	3	4
y	6	5	4	4	5	5	5	4	**3**

$$lev_{a,b}(i,j) = \left\{ \begin{array}{c} \max(i,j) \\ \min\left\{ \begin{array}{c} lev_{a,b}(i-1,j)+1 \\ lev_{a,b}(i,j-1)+1 \\ lev_{a,b}(i-1,j-1)+1_{a_i \neq b_j} \end{array} \right\} \end{array} \right\} \tag{1}$$

where 1 (a(i) ≠ b(j)) is the indicator function equal to 0 when a(i) = b(j) and equal to 1 otherwise, and lev(a, b) of (i, j) is the distance between the first i characters of a and the first j characters of b. Examples of the resulting matrix is shown in Table 2. The invariant maintained throughout the algorithm is that we can transform the initial segment s[1…i] into t[1…j] using a minimum of d[i, j] operations. Bottom-right element of the array contains the answer (distance between words "Saturday" and "Sunday").

3 Model Development

3.1 Model Architecture and Mechanics

The pairing system for a given address entered by free entry into the following fields: delivery line, domestic number of the recipient, postal code of the recipient, location of the recipient. As result, model is matching most probable address from the metadata repository, which best suits her and receives the relevant geo-coordinate (Fig. 2).

The sequence of operations in the system is as follows:

Step 1: In the first step, from the source system to the analytical platform filled with a reference address. The dynamics of the address base depends on the change of reference addresses;
Step 2: In the second step, the data on consignments from the source system, which combines shipments from several operating systems;
Step 3: In the third step, the pairing module retrieves the currently unpublished shipments;

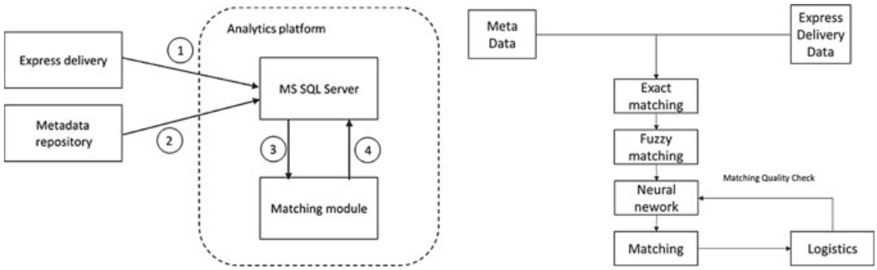


Fig. 2 Model architecture (mechanics on the left and matching module on the right)

Step 4: Finally, the pairing module performs the processing, and the pairing results return back to the same database.

3.2 Address Matching Workflow

The pairing module retrieves undelivered addresses from the central records, pairs them with the reference addresses in the address bar, and stores the results together with the geocache coordinates. The pairing process is shown in Fig. 3.

Workflow can be explained as follows: the system matches the addresses that fully match the names or aliases of the street and the place; addresses that are not paired in the previous step are paired with fuzzy matching, where couples above a certain degree of reliability are considered paired; pairs that are not paired in the previous

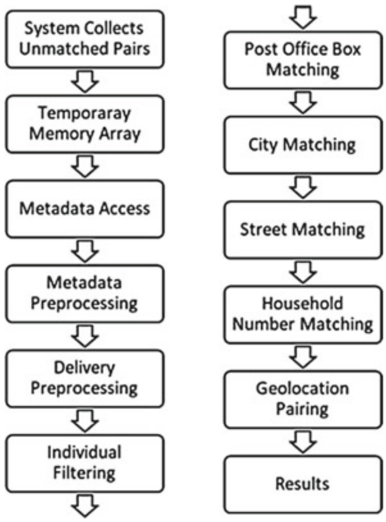


Fig. 3 Address matching workflow

steps are paired by a neural network that is initially trained on the basis of fuzzy matching and is continuously trained based on feedback from the logistics system.

4 Results

4.1 Address Tokenization/Decomposition

Electronic text is a linear sequence of symbols (characters or words or phrases). Before any real text processing is to be done, text needs to be segmented into linguistic units such as words, punctuation, numbers, alpha-numeric etc. This process is called tokenization. Tokenization is a kind of pre-processing in a sense; an identification of basic units to be processed. It is conventional to concentrate on pure analysis or generation while taking basic units for granted. Yet without these basic units clearly segregated it is impossible to carry out any analysis or generation. The identification of units that do not need to be further decomposed for subsequent processing is an extremely important one. Errors made at this stage are very likely to induce more errors at later stages of text processing and are therefore of great importance. To address this problem, a number of advanced methods which deal with specific challenges in tokenization have been developed to complement standard tokenizers. Tokenization is generally considered as a relative to other tasks in natural language, and one of the more uninteresting tasks. Due to fact addresses are already separated into “tokens”: city, street name, household number errors made in this phase are to be of lower importance however there is still need to understand common issues that can be propagated into later phases and cause problems. The process of segmenting running text into words and sentences needs to be robust in terms of combining for example “street name and number” or “city and street name” as one token. Most common error will be combination of “street name and household number” due to general way how users are careful when they input/edit data.

4.2 Address Matching, Confidence and Accuracy

As explained, model is using hierarchical structure to improve accuracy (from city, street name, household number towards back propagation using knowledge base from previous matching processes). As a combination being most accurate, I would like to point out usability of Levenshtein method in such environment. Each level was forced with Levenshtein distance measurement and verified through knowledge database.

Table 3 Address matching: city

City name: delivery	Matching model	Matching confidence	City name: match (metadata)
Tar (Torre)	4.0	65%	Staro Topolje
Barbariga	4.0	67%	Barban
Vrsar (Orsera)	2.0	69%	Vrsar
Laslovo (Szentlászló)	4.0	70%	Laslovo
Vodnjan (Dignano)	2.0	70%	Vodnjan
Višnjan (Visignano)	4.0	70%	Višnjan
Kaštelir (Castelliere)	4.0	73%	Kaštelir
Berek	1.0	73%	Berak
Brtonigla (Verteneglio)	4.0	75%	Brtonigla
Končanica (Končenice)	4.0	75%	Končanica
…	…	…	…
Galižana (Gallesano)	3.0	100%	Galižana
Umag (Umago)	3.0	100%	Umag
Selnica	3.0	100%	Selnica Šćitarjevska
Zemunik	3.0	100%	Zemunik Donji
Plavnice Gornje	3.0	100%	Gornje Plavnice
Zaton Veliki	3.0	100%	Zaton
Zagreb-Novi Zagreb	3.0	100%	Zagreb
Doli	1.0	100%	Zaton Doli
Ivanovo Selo*	3.0	100%	Ivanovo Selo

4.2.1 Address Matching: City

As for matches, top/bottom scoring (matching confidence) results are noted in Table 3. As it may be visually compared, for this particular purpose system was very accurate even with low detail string description on individual character level.

Visualization matrix showed matching confidence level on random sample is shown in Fig. 4. Methodology confirm very robust results having average confidence level on 80+% most of the time. Corresponding matrix represents different city name length where target lengths 15+ characters show rapidly decreased confidence level however still over 70%.

Same sample was challenged also on number of records showing very efficient confident level even on larger scale data comparing to sample (Fig. 5).

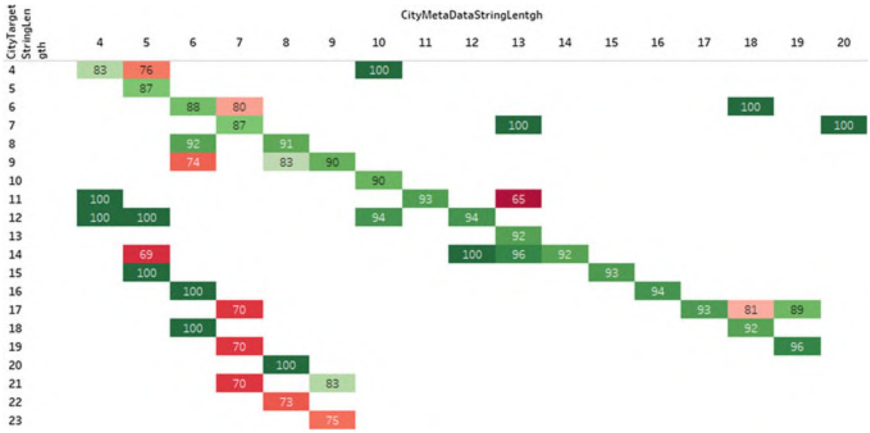


Fig. 4 City match string length versus confidence matrix

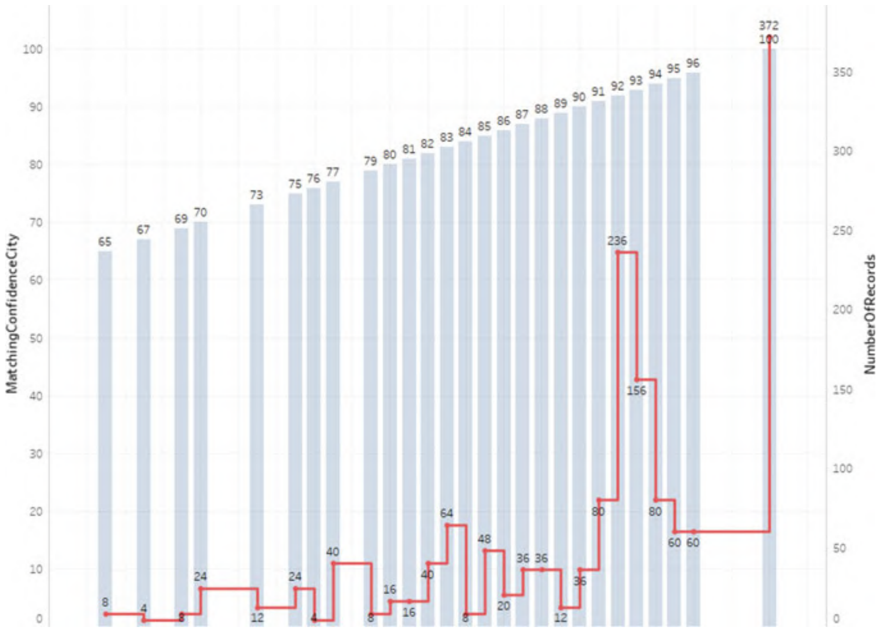


Fig. 5 City match confidence matrix versus number of records paired (sample)

4.2.2 Address Matching: Street Name

As for street name, our model showed even better results with street names being either expanded/shortened or strings were contextually “dirty” in ways familiar to address notation but unsuitable for machine learning (Table 4 and Fig. 6).

Table 4 Address matching: street name

Street name: delivery	Matching model	Matching confidence	Street name: match (metadata)
GAJEVA 39	2.0	50%	GAJEVA
BOZICKOVICEVA 8	3.0	53%	ULICA BISKUPA BOŽIČKOVIĆA
TRŽNICA MURTER	2.0	53%	MURTERSKIH ZADRUGARA
LIMSKA 8	2.0	55%	LIMSKA ULICA
DOMOVINSKOG RATA	2.0	56%	OMLADINSKO ŠETALIŠTE
STROSSMAYEROVA 18 A	2.0	56%	JOSIPA JURJA ŠTROSMAJERA
MALI IŽ	3.0	71%	VELI IŽ
MATTEO BENUSSI CIO 2	2.0	72%	MATTEA BENUSSIJA
PRILAZ STARE VOŠTARNICE	2.0	72%	STARA VOŠTARNICA
…	…	…	…
PUT SV ANTE	4.0	73%	PUT SVETOG ANTE PADOVANSKOG
I.G. KOVAČIĆA 9	2.0	73%	ULICA IVANA GORANA KOVAČIĆA
NAŠICKA	2.0	73%	NAŠIČKA
PUT BAGATA	4.0	73%	PUT BLATA
OSJECKA	2.0	73%	OSJEČKA
KLAICEVA	2.0	73%	KLANAC
KIŠPATIĆEVA	4.0	73%	ULICA MIJE KIŠPATIĆA
ZEMUNIK GORNJI 127	3.0	73%	DONJI ZEMUNIK ULICA I
MARKA VUCKOVICA	2.0	87%	MARKA VUČKOVIĆA
SV. BARBARE	2.0	87%	SVETE BARBARE

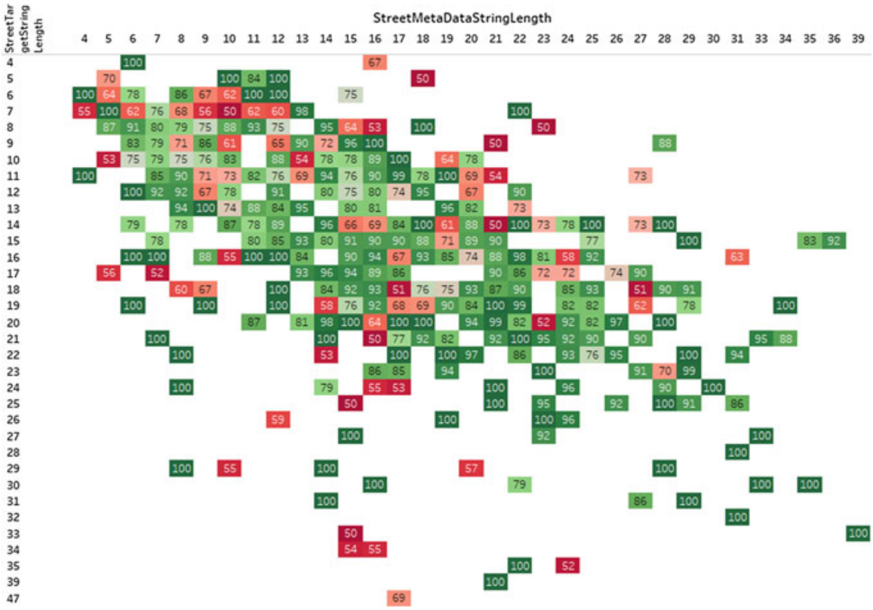


Fig. 6 Street name match string length versus confidence matrix

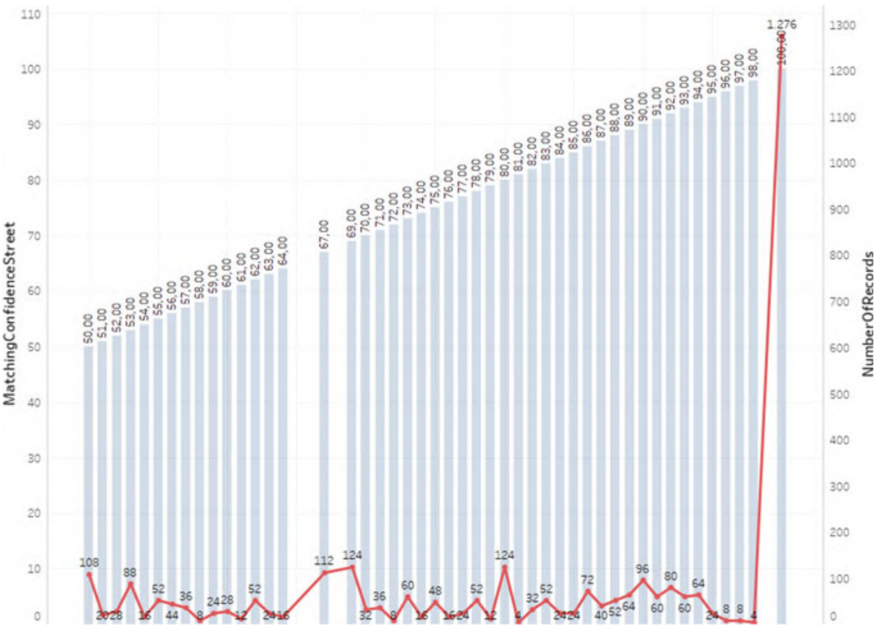


Fig. 7 Street name confidence matrix vs number of records paired (sample)

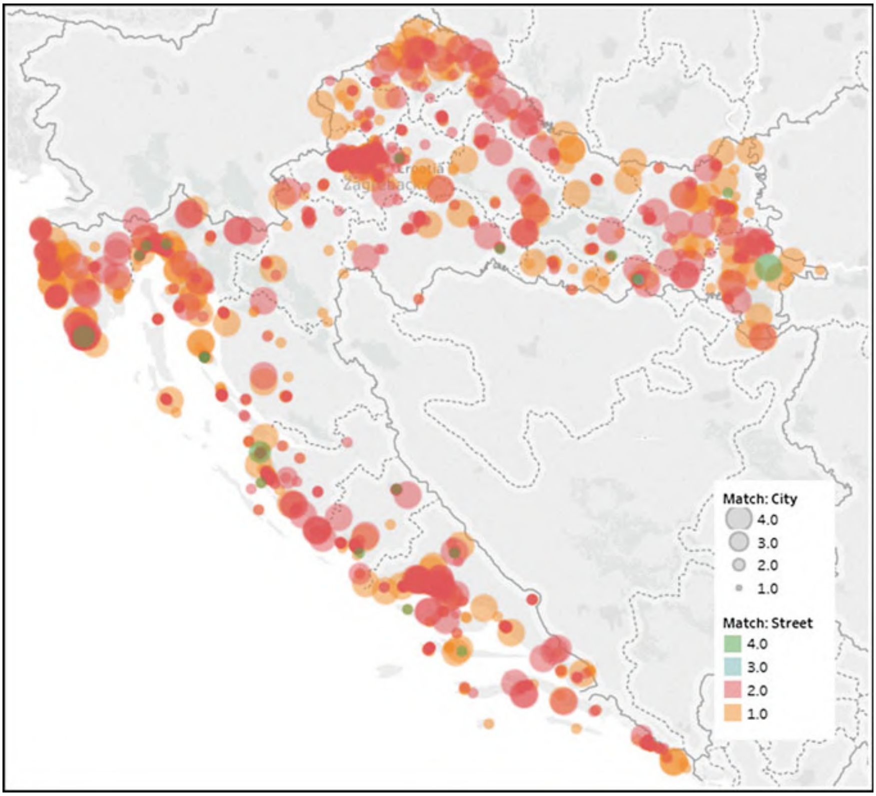


Fig. 8 Advanced model results visualization using geospatial data

Visualization matrix showed matching confidence level on random sample is shown below. Methodology still confirmed very robust results having average confidence level on 50+% most of the time.

Same sample was challenged also on number of records showing lower than city mane but still quite efficient confident level even on larger scale data comparing to sample (Fig. 7).

Similar approach was used for household number matching. When combining all matching levels together, our methodology stands up as robust and very useful for this specific purpose.

4.2.3 Address Matching: Advanced Visualization Using Geospatial Data

With geolocation data being central logistics component for delivery process, matching system can be visualized using common tools creating powerful dashboard for policy makers and overall system management purposes. Example of visualization dashboard using sample data on map of Croatia is show in Fig. 8.

5 Conclusion

Research shown that string distance measurements can be very efficiently used as support in delivery services in logistic. Distance measurement NLP methods are widely usable in text mining and can be used to find the similarity among sentence or document. As part of logistics process, being able to determine correct address using machine learning we need to tackle issue of two addresses comparison (street name, city name etc.) is crucial for efficient service. This paper explains comparison techniques based on similarity score that can be calculated using distance measurement. As part of process, several distance measurements were compared while conclusion include results and recommendation on usage in address and location matching in logistics (post services). As final model, Levenshtein method was used in combination with original hierarchy bundled into robust address matching model with great result and future potential.

References

1. Navarro, G.: A guided tour to approximate string matching (PDF). ACM Comput. Surv. **33**(1), 31–88 (2001). CiteSeerX 10.1.1.452.6317. https://doi.org/10.1145/375360.375365
2. Wagner, R.A., Fischer, M.J.: The string-to-string correction problem. J. ACM **21**(1), 168–173 (1974). https://doi.org/10.1145/321796.321811 (Author, F.: Article title. Journal 2(5), 99–110 (2016))
3. Backurs, A., Indyk, P.: Edit distance cannot be computed in strongly subquadratic time (unless SETH is false). In: Forty-Seventh Annual ACM on Symposium on Theory of Computing (STOC) (2015). arXiv:1412.0348. Bibcode:2014arXiv1412.0348B
4. Abrahamson, K.: Generalized string matching. SIAM J. Comput. **16**(6), 1039–1051 (1987). https://doi.org/10.1137/0216067
5. Alstrup, S., Brodal, G.S., Rauhe, T.: Pattern matching in dynamic texts. In: Proceedings of the Eleventh Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 819–828, 09–11 Jan 2000, San Francisco, California, USA (2000)
6. Andoni, A., Krauthgamer, R., Onak, K.: Polylogarithmic approximation for edit distance and the asymmetric query complexity. IEEE Symp. Foundations of Computer Science (FOCS) (2010). arXiv:1005.4033. Bibcode:2010arXiv1005.4033A. CiteSeerX 10.1.1.208.2079
7. Levenshtein, I.: Binary codes capable of correcting deletions, insertions and reversals. Doklady Akademii Nauk SSSR **163**(4), 845–848 (1965) also Soviet Physics Doklady **10**(8) 707–710 (1966)
8. Needleman, S.B., Wunsch, C.D.: A general method applicable to the search for similarities in the amino acid sequence of two proteins. J. Mol. Biol. **48**, p443–p453 (1970)
9. Anderson, R.J., Miller, G.L.: Deterministic parallel list ranking. Algorithmica **6**, 859–868 (1991)
10. Cole, R., Hariharan, R.: Approximate string matching: a simpler faster algorithm. In: Proceedings of the Ninth Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 463–472, 25-27 Jan 1998, San Francisco, California, USA (1998)

11. Henzinger, M., Raghavan, P., Rajagopalan, S.: Computing on data streams. Tech. Rep. SRC 1998-011, DEC Systems Research Centre (1998)
12. Crochemore, M., Rytter, W.: Text Algorithms. Oxford University Press Inc., New York (1994)
13. Goel, A., Indyk, P., Varadarajan, K.: Reductions among high dimensional proximity problems. In: Proceedings of the Twelfth Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 769–778, 07–09 Jan 2001, Washington, DC, USA (2001)
14. Goldberg, A., Plotkin, S., Shannon, G.: Parallel symmetry-breaking in sparse graphs. In: Proceedings of the Nineteenth Annual ACM Symposium on Theory of Computing, pp. 315–324, Jan 1987, New York, New York, USA (1987). https://doi.org/10.1145/28395.28429
15. Gusfiel, Dan: Algorithms on Strings, Trees, and Sequences: Computer Science and Computational Biology. Cambridge University Press, New York (1997)
16. Indyk, P.: Stable distributions, pseudorandom generators, embeddings and data stream computation. In: Proceedings of the 41st Annual Symposium on Foundations of Computer Science, p. 189, 12–14 Nov 2000 (2000)
17. Karp, R.M., Rabin, M.O.: Efficient randomized pattern-matching algorithms. IBM J. Res. Dev. **31**(2), 249–260 (1987). https://doi.org/10.1147/rd.312.0249
18. Landau, G.M., Vishkin, U.: Efficient string matching with k mismatches. Theor. Comput. Sci. **43**(2-3), 239–249 (1986)
19. Levenshtein, V.I.: Binary codes capable of correcting deletions, insertions and reversals. Sov. Phys. Dokl. **10**(8), 707–710 (1966)
20. Sankoff, D., Kruskal, J.B.: Time Warps, String Edits and Macromolecules: The Theory and Practice of Sequence Comparison. Addison-Wesley (1983)
21. Landau, G.M., Vishkin, U.: Fast parallel and serial approximate string matching. J. Algorithms **10**(2), 157–169 (1989) https://doi.org/10.1016/0196-6774(89)90010-2
22. Masek, W.J., Paterson, M.S.: A faster algorithm computing string edit distances. J. Comput. Syst. Sci. **20**, 18–31 (1980)
23. Shapira, D., Storer, J.A.: Edit distance with move operations. In: Proceedings of the 13th Annual Symposium on Combinatorial Pattern Matching, pp. 85–98, 03–05 July 2002 (2002)
24. Mrsic, L., Skansi, S., Kopal, R.: Preliminary study for a survey-based fuzzy membership function definition for imprecise quantification in Croatia. In: Central European Conference on Information and Intelligent Systems (2018). ISSN 1847-2001, e-ISSN 1848-2295
25. Mrsic, L., Kopal, R., Klepac, G.: Analyzing slavic textual sentiment using deep convolutional neural networks. In: Intelligent Decision Support Systems for Sustainable Computing: Paradigms and Applications, vol. 705, pp. 207–224 (2017). https://doi.org/10.1007/978-3-319-53153-3_11

An Investigation of Information Granulation Techniques in Cybersecurity

Sani Suleiman Isah, Ali Selamat, Roliana Ibrahim and Ondrej Krejcar

Abstract Information Granulation in the context of Granular Computing provides a viable alternative for finding solutions to complex problems using granules. Issues such as intrusions, malware exigencies, spam and user unauthorized access still remain challenging in cybersecurity. Moreover, as the prevalence of undetected attacks due to system design flaws and system development flaws become rampant in the cybersecurity systems. Although, numerous techniques that have been applied have shown very good prospects, there are several difficulties in managing cyber-attacks from the angle of biometric recognition systems which are commonly used in cybersecurity. These challenges has positioned cybersecurity issues to be regarded as complex and uncertain which requires techniques such as information granulation to unravel a sustainable solution. This paper investigates how information granulation techniques are used in cybersecurity detection models with the aim of providing a holistic view of the current status of research in this area. In this paper, we proposed a framework that applied the principle of justifiable granularity (PJG) in the feature extraction module of a finger-vein recognition system using granular support vector machines as classifier to justify the effectiveness of information granulation in

S. S. Isah · A. Selamat (✉) · R. Ibrahim
Faculty of Engineering, School of Computing, UTM & Media and Games Center of Excellence (MagicX), Universiti Teknologi Malaysia, Johor Bahru, Malaysia
e-mail: aselamat@utm.my

S. S. Isah
e-mail: ssmaiadua@gmail.com

R. Ibrahim
e-mail: roliana@utm.my

A. Selamat
Malaysia Japan International Institute of Technology (MJIIT), Universiti Teknologi Malaysia, Jalan Sultan Yahya Petra, Kuala Lumpur, Malaysia

A. Selamat · O. Krejcar
University of Hradec Kralove, Rokitanskeho 62, 500 03 Hradec Králové, Czech Republic
e-mail: ondrej.krejcar@uhk.cz

M. Huk et al. (eds.), *Intelligent Information and Database Systems: Recent Developments*, Studies in Computational Intelligence 830,
https://doi.org/10.1007/978-3-030-14132-5_12

strengthening a verification system in a cybersecurity setting. We benchmark our result with state-of-the-art biometric verification systems, and our approach shows promising contribution in that direction.

Keywords Information granulation · Cybersecurity · Biometrics · Verification · Granular support vector machines

1 Introduction

Information granulation as originally introduced by Zadeh [1] provides a viable alternative in finding solution to complex problems. As posited by [1], when a problem is regarded as vague, uncertain and imprecise, it is best described with the use of granules. And granules are created using an information granulation technique [1, 2] and it results to information granule(s) (IGs). A created IG is characterized as being a cluster, subset or object that composed of similar, functional, and indistinguishable granule within a close proximity. And a granule can be created using several formalism of information granulation, such as fuzzy sets, rough sets, set intervals, shadowed sets to mention a few. Information granulation has two major approaches in constructing namely; top-down and bottom-up information granulation approaches, the summary can be found in Table 1.

Furthermore, information granulation (creation, representation and knowledge discovery of granules) still remain a crucial process [4, 5], which has found significant level of contribution in many areas such as system modeling, artificial intelligence, data mining and cybersecurity. Information granulation provides a high level view of the data by unraveling hidden and uncertain features or attributes that were somewhat not seen with traditional modeling or description techniques. One of the major issue with information granulation procedure, is the difficulty in determining an "optimal information granule" [4], an information granule that is efficient and accurate. Nonetheless, researchers have used many of the available methods or formalism to construct an information granule in isolation such as rough [7], fuzzy [8] or in combination such as fuzzy-rough [9] with some optimization techniques

Table 1 Information granulation approaches

S/N	Paper	Problem	Approach	Result	Limitation
1	[3]	Intrusion detection	Top-down	Rough sets	High dimensional data
2	[4]	Model interpretation	Bottom-up	Fuzzy sets	Large dataset
3	[5]	Decision making	Top-down	Set intervals	Linguistic data
4	[6]	Design of IG	Top-down	Weighted interval	Spatial data

utilized. Also Pedrycz and Homenda [10] proposed the principle of justifiable granularity (PJG), which advocates creation of an optimal information granule through coverage (experimental evidence) and specificity (semantics). The idea of this principle is that a created information granule must accommodate as many experimental data as possible (coverage) as well as be as specific as possible in order to have semantic meaning. The work of [10] has gained attention of researchers towards its well-established mathematical approach [5–7].

It is worth noting that information granulation techniques are schematically divided into two; principled-based and formalism-based. Principled-based are the type of information granulation that follow the major granulation principles such as the principle of justifiable granularity (PJG) [10], which advocates creation of an optimal information granule through coverage (experimental evidence) and specificity (semantics). On the other hand, formalism-based approach are those that employ the direct use of reasoning formalism such as fuzzy sets, rough sets, set intervals, shadowed sets and many others in which these two major approaches in constructing information granules could be followed. Either the top-down or bottom-up information granulation approaches which covers different ways in which complex problems are tackled depending on type and complexity of a problem.

The use of information granulation-based solutions have permeated into several domains particularly cybersecurity [11]. In cybersecurity, we encounter various techniques of computational intelligence such as biometric recognition systems aiming at neutralizing the challenges in cyber crimes such as Intrusions, and user unauthorized access. Many researches have used information granulation techniques to tackle issues in cybersecurity domain [12–22] (summary in Table 2). In Yang et al. [23], recognition efficiency of finger-vein biometric system is improved using a hyper-sphere granular computing model. And their approach has performed well in improving recognition accuracy. However, tackling prevalence of undetected attacks due to system design flaws and system development flaws [24] has positioned cybersecurity to be complex and uncertain which requires techniques such as information granulation to unravel a sustainable solution.

In this direction, this paper investigates the application of information granulation techniques in cybersecurity through biometric pattern recognition and algorithmic strength. While there have been a number of studies devoted to information granulation techniques [23], the originality of the investigation of this study is how the use of information granulation techniques could provide a viable defence mechanisms in cybersecurity through biometric pattern recognition. Therefore, by taking the advantage of PJG [10], in this study, we applied the PJG in finger-vein biometric recognition system using granular support vector machines (GSVM) as classifier to improve recognition accuracy thereby justifying the effectiveness of information granulation in strengthening a verification system in cybersecurity settings. The paper is organized as follows: Information granulation and different approaches to information granulation used in cybersecurity is presented in Sect. 2 under related

Table 2 Information granulation techniques used in cybersecurity systems

Paper	Method	Cyber-system	EER (%)	Accur (%)	I.G Technique	Limitation
[13]	Similarity fusion	Biometric	0.135	98	✓	Classification
[14]	Neuro-fuzzy	Power	x	97	✓	Accuracy
[15]	ANFIS and SVM	Medical data	x	81.90	✓	Classification
[16]	Neighborhood sets	Biometeric	x	70.2	✓	F.Extraction
[12]	PJG	Critical infr.	x	x	✓	Classification
[18]	Fourier and fuzzy	Biometric	4.82	x	✓	Accuracy
[19]	Interval-fuzzy	Biometric	x	96	✓	Segmentation
[20]	Visual analysis	Biometric	5	x	✓	Optimization
[21]	Hierachical model	Biometric	x	99.78	✓	F.Extraction
[22]	Fuzzy sets	Biometric	6.00	x	✓	Applicability

works. Sect. 3 presents the proposed approach. The experimental results of the proposed approach in the use of PJG with GSVM to improve verification accuracy is presented in Sect. 4. Section 5 is the conclusion of the study.

2 Related Works

Different approaches in the use of information granulation has made the technique widely implemented in research activities, especially in the challenging domain of cybersecurity where firewalls, encryption and access control fail to completely protect the systems and networks. The prevention of unauthorized access could be a serious challenge since the data and information presented to cybersecurity systems is permeated with uncertainty. Many researchers have used information granulation technique to solve complex problems from the perspective of uncertainty, vagueness and impreciseness. In this section, related works and motivation to our work is presented.

In [9] a rough cognitive network granular model is used to tackle intrusion detection which was able to detect abnormal patterns in the network traffic. In Bay and Yang [25], they applied a pixel-based fusion method for finger recognition to create an atomic hypersphere granules with fuzzy inclusion measure to compute similarity between two fusion hypersphere granules in which provides a reliable image matching.

In Fujita et. al [12] the challenging resilience in critical infrastructure was investigated. The authors, in their quest for approximate reasoning and rapid decision making due to uncertainty involved in attacks and natural hazards, create a granular structures based on binary relations to provide inter-dependency in critical infrastructure. Their approach, provide an enhanced reasoning and modeling capability on intentional attacks to a critical infrastructure. However, their approach lacks optimization of created granular structure interval bounds in order to maximize coverage due to multi-dependency paths which could be a vulnerability for attack.

In Sareen et. al [26], information granulation technique using fuzzy k-nearest neighbor was combined with secret sharing scheme in a two stage process to prevent unauthorized access and protect confidentiality of sensitive information for remote users of mobile medical applications. But in this work no evaluation on the technique used and also due to no optimization procedure involved, the specificity and coverage [10] are in serious conflict and may lead to vulnerability of the defence mechanism.

In Salehi et. al [27], a granular classifier was designed by hybridizing particle swarm optimization and k-means algorithm in order to solve a pattern recognition problem in spam detection. The discovered granules were optimized to cover noise points and patterns in the feature space in order to enhance the performance of the primary structure. However, their work could not provide an optimal information granule that can cover the global parameters in order to expand the discovery of the granules and noisy points.

A multimodal finger-based recognition method based on information granulation was proposed by Li et. al [13]. In this work, a bottom-up and top down information granulation approaches were used interchangeably to represent finger features effectively with a higher accuracy recognition rate. Which also from the two approaches, they found that fine-granularity space outperforms coarse-granularity space with a higher matching accuracy while coarse-granularity has higher matching efficiency. However, issue of overlapping remain unresolved.

From the above outlined contributions, we can deduct how several areas of cybersecurity ranging from intrusion detection [9], access control [9, 25], critical infrastructure [12], and spam detection [27] have been confronted using information granulation techniques. However, the prevalence of undetected attacks through impostors and unauthorized access as a result of flaws in the design and implementation of cybersecurity systems remain unresolved. We intend to fill this gap by taking biometric recognition system as case study to hybridized information granulation with discriminative classifier (granular support vector machines) in a finger-vein classification. We used finger-vein dataset from University Sains Malaysia (USM) (FV-USM), and in the experiments, our classification approach has shown an improved learning process and high accuracy towards classification of genuine and imposters in the dataset [17]. We believe this is facilitated in the use of information granulation technique before the feature extraction stage of the authentication process (see Fig. 1).

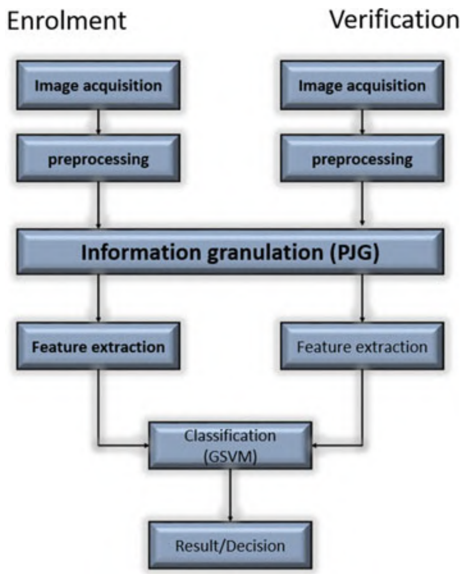


Fig. 1 Proposed approach to improve Cybersecurity biometric system

3 Proposed Approach

In this section, we describe the proposed approach which facilitate learning process in the detection of unauthorized access and identification of imposters. Generally, this approach, would help cybersecurity system designers and implementers in overcoming the prevalence of undetected attacks. We start by discussing some preliminary concepts relevant in the design of cybersecurity systems.

3.1 Preliminaries

Definition of Cybersecurity:- Cybersecurity is an umbrella term for high level approach to the security of data and provision of secured access to information through secured infrastructure [28]. Cybersecurity play a major role in the protection from intentional threats and unauthorized access. In recent years adoption of cybersecurity has changed organizational in-house legacy information security strategies to the cloud-based due to over reliance on Internet of Things (IoT). The challenges in cybersecurity in dealing with vulnerabilities before they manifest requires tools that will help in getting intelligence about threats from large datasets of variable capabilities such as big data [29]. These vulnerabilities, raised concern by researchers and practitioners to provide attempted solutions from artificial intelligence, machine learning, deep learning and big data analytics perspectives [29].

Existing cybersecurity systems can be aligned to biometric-based identification and verification systems through protection of unauthorized access and provision of secured and reliable access to genuine users. Such systems covers technologies such as pattern recognition, computer vision, and image processing. A state-of-the-art cybersecurity system can be envisioned in image processing [30], where physical intrusions and security bypass could be challenged through enhanced biometric-based and verification systems. The major mechanisms are irises, faces, fingerprints, finger-veins etc. All of these methods have prospects and limitations, and overcoming their challenges has been an active area by researchers. In the following subsections, we would discuss on the biometrics in cybersecurity systems, vulnerabilities in cybersecurity systems, and threats/attacks commonly in cybersecurity systems perspectives.

Biometrics in Cybersecurity Systems:- Cybersecurity is referred to as security of control systems [28]. Cybersecurity systems could be categorized into Computer systems and network systems. Under computer systems, we acknowledge the data-information security, the server-side security and application-based security. In the network systems, we encounter critical infrastructures and hardware devices dedicated for communication. Several cybersecurity systems have been developed over the years in relation to the control of cyber threats especially towards to the prevention of intrusions and unauthorized access [30]. An example of such systems is the biometric identification and verification systems that has been in the use of human Physical traits such as; voice, face, fingerprints, iris and finger-vein to provide a detection mechanism of unauthorized access to cybersecurity systems. Among these traits, the finger-vein biometric recognition system have been regarded as accurate and fraud-proof biometric technique. Although the finger-vein may be disadvantageous due to unreliable and resilient mechanisms that will accommodate the degradation of the images due to light attenuation in biological tissue upon capture, low image contrast, vein thickness and noise existence [31]. Moreover, most feature extraction algorithms used in biometric authentication takes longer time for high dimensional data [31], causing high matching time and excessive training of sample data.

However with approaches such as information granulation techniques, the highlighted problems would be highly improved. In Yang et al. [32] a granular hierarchical hyper-sphere model was designed to improve the finger-vein recognition by enhancing the venous region and providing an improved handling of region of interest (ROI) extraction. Even though their work is significant towards improving the finger-vein biometric recognition, it still needed a better feature generation method than the traditional PCA they used which has strictly relies on linearly correlated data. These limitations of feature extraction methods, motivated the strength of our proposed model where the information granulation plays an important role after preprocessing of the input data.

Vulnerabilities:- Vulnerabilities refer to areas of soft target by attacks and security bypasses. Control systems can become vulnerable due to flaws in system design and development [29]. There will always be a vulnerability in any system, and threats tend to target and exploit related vulnerabilities and have a negative impact to the cybersecurity system. However, researchers have developed strategies to mitigate these vulnerabilities. In [33], the authors proposed a mathematical analysis as a process to identify vulnerabilities before they become targets for attackers. The authors argue that most vulnerabilities are as a result of poor software design and implementations that could give way for attacks such as denial of service, remote access, code injections and privilege escalation. In [28], a quantitative vulnerability assessment model was proposed in order to analyze potential physical consequences of cyber-attacks on a smart grid distribution automation system. The model was designed based on a two-level process of terminal device level and control center server level in order to simulate the attack process.

Threats and Attacks:- Several types of attacks have been recorded over the years on cybersecurity systems. Which comprises of phishing, sql injections, cross-site scripting, man-in-the-middle, malware, denial of service and Speer phishing all in the name of getting access to security systems. Currently, active attacks on biometrics are targeted at the vulnerabilities associated with the design and implementations. To reduce these threats and attacks, cancelable algorithms [31] are designed to distort an image and record the distortions into generated templates. Therefore, if an image is compromised as a result of attack, it can easily be changed since the original image was not saved anywhere after the distortion. Now the issue is the ability to design protected algorithms, even though no foolproof but enhanced and optimized. Information granulation techniques could be a promising solution to the protection of biometric cancelable algorithms by enhancing the feature generation methods through hierarchical attributes of granular computing [32].

3.2 Proposed Method

The application of information granulation in cybersecurity systems development would provide a synergistic control of vulnerabilities and potential attacks, because of the relationship in terms of uncertainty, vagueness and imprecision of cyber-attacks. It has been stated in the literature that most existing attacks target vulnerabilities in weak or flawed system designs [34]. Therefore we propose a framework that will guide the development of cybersecuirty systems at a higher level. The framework will allow the integration of information granulation techniques into the development of cybersecurity systems at all levels, especially the feature extraction level where increase in efficiency, calculating speed, data legibility and classification performance is guaranteed [14]. This will drastically reduce flaws and weaknesses being experienced in most current designs. We consider one of the biometric systems as a case study to justify the effectiveness of the integration of information granulation techniques in the design process. Moreover, the formalism adopted for the information granulation is the principle of justifiable granularity [10]. And since the adopted principle is generalized in the sense that it can use any reasoning concept, we employ fuzzy information granulation [1] to obtain segmentation boundaries. In segmentation, finding threshold is very important which fuzzy information granulation is efficient for that. Therefore the use if information granulation is to as well improve the segmentation results. Furthermore, after the granular features are extracted, we employ granular support vector machines (GSVM) [35] in order to discriminate between genuine and imposter images in the dataset. The proposed framework is presented in Fig. 1.

A viable and sustainable system from cybersecurity perspective must as a matter of importance consider the three most important factors; uncertainty, vagueness and imprecision. These factors are related to unknown data that will be processed during the system implementation, therefore information granulation can provide the detail wanted or hide the unwanted detail for optimal realization of suitable outcome. We present the basic mathematical concept of our approach.

4 Experimental Results

In this section we present the experimental results based on the database of finger-vein from University Sains Malaysia (FV-USM) [17]. The database contains 492 fingers with 12 images each making 5,904 images. The images are 8-bit grey scale level of JPG files with resolution of 640×480 pixels in raw images. The region of interest (ROI) has already been extracted by the database authors. The experiments are implemented in MATLAB2017b on a PC with Intel® Core^{TM} i7-3770 CPU @ 3.40 GHz 3.40 GHz installed with 16 GB RAM. All images are are normalized and scaled to 120×60 pixels. The list of information granulation techniques used in the design of cybersecurity systems such as biometrics is listed in Table 2.

As stated earlier, the images acquired, have their ROI extracted already by the authors. We applied image enhancement techniques using contrast limited adaptive histogram equalization (CLAHE) (see Fig. 2.). And we further perform fuzzy information granulation using principle of justifiable granularity (PJG) by formulation two objective optimization problem based on the coverage data and specificity. Information granules are created using fuzzy c-means clustering [36]. And furthermore, we perform feature extraction using local binary pattern [37], and we classify the data using the granular support vector machines which the classification result is presented in Table 3. In order to show the superiority of our approach in terms of classification and recognition performance, we repeat the same experiment with other formalisms used in finger-vein biometric authentication systems in the same order. The results in terms of EER are presented in Table 3. From the results, it is observed that the performance of of the proposed method is significantly better than other variants. The recognition accuracy rate of our approach is slightly better than the neighbourhood sets counterpart. We believe that the fuzzy information granulation can provide better segmentation to aid the local binary pattern extraction. It should also be noted, that due to limited space images obtained in all experiments are reserved for the conference presentations.

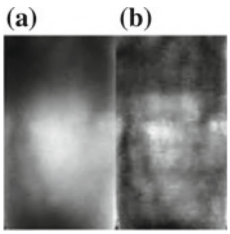


Fig. 2 Image enhancement using CLAHE **a** before **b** after enhancement

Table 3 Classification results with comparison of approaches based on recognition rate and EER

Formalism	Recognition rate(%)	EER(%)	Classification time(s)
Rough set	89.04	3.46	0.52 ± 0.01
Set neighborhood	92.40	2.89	0.45 ± 0.09
Probability systems	82.12	3.12	0.261 ± 0.27
Fuzzy PJG (our proposed)	94.20	2.48	0.22 ± 0.03

5 Conclusion

This paper combine the concept of cybersecurity and granular computing to provide a viable solution in the design of intelligent systems. Biometric recognition systems are intelligent systems that play an important role in cybersecurity for the protection of valuable information and properties. The important role of biometric systems towards providing security to privacy and authorized user access justify the use as finger-vein biometric system as our case study. Although, this study is a preliminary study towards the development of a robust finger-vein biometric recognition with improved image degradation and an enhanced extraction method for region of interest (ROI) and segmentation using information granulation approach. Definitely, the anticipated solution would guarantee a more scalable cyber activities in relation to biometric identification and verification systems.

Acknowledgements The authors wish to thank Universiti Teknologi Malaysia (UTM) under Research University Grant Vot-20H04 supported under Ministry of Higher Education Malaysia for the completion of the research.

References

1. Zadeh, L.A.: Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic. Fuzzy Sets Syst. **90**, 111–127 (1997). https://doi.org/10.1016/S0165-0114(97)00077-8
2. Yao, Y.: A triarchic theory of granular computing. Granul. Comput. **1**, 145–157 (2016). https://doi.org/10.1007/s41066-015-0011-0

3. Nápoles, G., Grau, I., Falcon, R., Bello, R., Vanhoof, K.: A granular intrusion detection system using rough cognitive networks. In: Abielmona, R., Falcon, R., Zincir-Heywood, N., Abbass, H.A. (eds.), Recent Advances in Computational Intelligence in Defense and Security, pp. 169–191. Springer International Publishing, Cham (2016). https://doi.org/10.1007/978-3-319-26450-9-7
4. Zhu, X., Pedrycz, W., Li, Z.: Granular data description: designing ellipsoidal information granules. IEEE Trans. Cybern. **47**, 4475–4484 (2017). https://ieeexplore.ieee.org/abstract/document/7589082
5. Cabrerizo, F.J., Morente-Molinera, J.A., Pedrycz, W., Taghavi, A., Herrera-Viedma, E.: Granulating linguistic information in decision making under consensus and consistency. Expert Syst. Appl. **99**, 83–92 (2018). https://doi.org/10.1016/j.eswa.2018.01.030
6. Wang, D., Pedrycz, W., Li, Z.: Granular data aggregation: an adaptive principle of the justifiable granularity approach. IEEE Trans. Cybern. 1–10 (2018). https://doi.org/10.1109/TCYB.2017.2774831
7. Liu, S., Pedrycz, W., Gacek, A., Dai, Y.: Development of information granules of higher type and their applications to granular models of time series. Eng. Appl. Artif. Intell. **71**, 60–72 (2018). https://doi.org/10.1016/j.engappai.2018.02.012
8. Zhu, X., Pedrycz, W., Li, Z.: A design of granular Takagi-Sugeno fuzzy model through the synergy of fuzzy subspace clustering and optimal allocation of information granularity. IEEE Trans. Fuzzy Syst. **6706** (2018). https://doi.org/10.1109/TFUZZ.2018.2813314
9. Nápoles, G., Grau, I., Falcon, R., Bello, R., Vanhoof, K.: A granular intrusion detection system using rough cognitive networks. In: Abielmona, R., Falcon, R., Zincir-Heywood, N., Abbass, H.A. (eds.) Recent Advances in Computational Intelligence in Defense and Security, pp. 169–191. Springer International Publishing, Cham (2016). https://doi.org/10.1007/978-3-319-26450-9-7
10. Pedrycz, W., Homenda, W.: Building the fundamentals of granular computing: a principle of justifiable granularity. Appl. Soft Comput. J. **13**, 4209–4218 (2013). https://doi.org/10.1016/j.asoc.2013.06.017
11. lzak, D., Chadzyska-Krasowska, A., Holland, J., Synak, P., Glick, R., Perkowski, M.: Scalable cyber-security analytics with a new summary-based approximate query engine. In: Proceedings—2017 IEEE International Conference on Big Data, Big Data 2017, pp. 1840–1849, 2018–Jan (2018). https://doi.org/10.1109/BigData.2017.8258128
12. Fujita, H., Gaeta, A., Loia, V., Orciuoli, F.: Resilience analysis of critical infrastructures: a cognitive approach based on granular computing. IEEE Trans. Cybern. 1–14 (2018). https://doi.org/10.1109/TCYB.2018.2815178
13. Li, Y., Peng, J., Zhong, Z., Jia, G., & Yang, J.: A multimodal finger-based recognition method based on granular computing. In: Sun, Z., Shan, S., Sang, H., Zhou, J., Wang, Y., Yuan , W. (eds.) Biometric Recognition, pp. 458–464. Springer International Publishing, Cham (2014)
14. Sun, W., Zhang, J., Wang, R.: Predicting electrical power output by using granular computing based Neuro-Fuzzy modeling method. In: Proceedings of the 2015 27th Chinese Control Decision Conference CCDC 2015, pp. 2865–2870 (2015). https://doi.org/10.1109/CCDC.2015.7162415
15. Al-Shammaa, M., Abbod, M.F.: Granular computing approach for the design of medical data classification systems. In: 2015 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology CIBCB 2015, pp. 1–7 (2015). https://doi.org/10.1109/CIBCB.2015.7300328
16. Vimitha, K., Jayasree, M.: Recognizing faces from surgically altered face images using granular approach. In: Proceedings of the 2017 International Conference on Wireless Communications, Signal Processing and Networking, WiSPNET 2017, pp. 463–466. (2017). https://doi.org/10.1109/WiSPNET.2017.8299799
17. Mohd Asaari, M.S., Suandi, S.A., Rosdi, B.A.: Fusion of band limited phase only correlation and width centroid contour distance for finger based biometrics. Expert Syst. Appl. **41**(7), 3367–3382 (2014). https://doi.org/10.1016/j.eswa.2013.11.033

18. Alpar, O., Krejcar, O.: Online signature verification by spectrogram analysis. Appl. Intell. **48**(5), 1189–1199 (2018). https://doi.org/10.1007/s10489-017-1009-x
19. Alpar, O., Krejcar, O.: A comparative study on chrominance based methods in dorsal hand recognition: single image case. In: Mouhoub, M., Sadaoui, S., Ait Mohamed, O., Ali, M. (eds.) Recent Trends and Future Technology in Applied Intelligence, pp. 711–721. Springer International Publishing, Cham (2018)
20. Kolda, L., Krejcar, O., Selamat, A., Brida, P., Kuca, K.: Contactless identification system based on visual analysis of examined element. In: Sieminski, A., Kozierkiewicz, A., Nunez, M., Ha, Q.T. (eds.) Modern Approaches for Intelligent Information and Database Systems, pp. 409–420. Springer International Publishing, Cham (2018). https://doi.org/10.1007/978-3-319-76081-0-35
21. Yang, J., Wei, J., Shi, Y.: Accurate ROI localization and hierarchical hyper-sphere model for finger-vein recognition. Neurocomputing 1–11 (2018). https://doi.org/10.1016/j.neucom.2018.02.098
22. Kolda, L., Krejcar, O.: Biometrie hand vein estimation using bloodstream filtration and fuzzy e-means. In: IEEE International Conference on Fuzzy Systems, pp. 0–5 (2017). https://doi.org/10.1109/FUZZ-IEEE.2017.8015736
23. Yang, J., Yang, Y., Liu, Z., and Shi, Y.: Hierarchical structure construction based on hypersphere granulation for finger-vein recognition. In: Yang, J., Hu, Q., Cheng, M.-M., Wang, L., Liu, Q., Bai, X., Meng, D. (eds.) Computer Vision, pp. 375–386. Springer Singapore, Singapore (2017)
24. Knowles, W., Prince, D., Hutchison, D., Disso, J.F.P., Jones, K.: A survey of cyber security management in industrial control systems. Int. J. Crit. Infrastruct. Prot. **9**, 52–80 (2015). https://doi.org/10.1016/j.ijcip.2015.02.002
25. Bai, G., Yang, J.: A new pixel-based granular fusion method for finger recognition. In: Eighth International Conference on Digital Image Processing (ICDIP 2016). SPIE (2016)
26. Sareen, S., Gupta, S.K., Sood, S.K.: An intelligent and secure system for predicting and preventing Zika virus outbreak using Fog computing. Enterp. Inf. Syst. **11**, 1436–1456 (2017). https://doi.org/10.1080/17517575.2016.1277558
27. Salehi, S., Selamat, A., Krejcar, O., Kuca, K.: Fuzzy granular classifier approach for spam detection. In: Núñez, M., Nguyen, N.T., Camacho, D., Trawiński, B. (eds.) Computational Collective Intelligence, pp. 256–264. Springer International Publishing, Cham (2015)
28. Ye, X., Zhao, J., Zhang, Y., Wen, F.: Quantitative vulnerability assessment of cyber security for distribution automation systems. Energies **8**, 5266–5286 (2015). https://doi.org/10.3390/en8065266
29. Aslan, Ö., Samet, R.: Mitigating cyber security attacks by being aware of vulnerabilities and bugs. In: Proceedings—2017 International Conference on Cyberworlds, CW 2017—Coop. With Eurographics Association. International Federation for Information Processing. ACM SIGGRAPH, 2017–Janua, pp. 222–225 (2017). https://doi.org/10.1109/CW.2017.22
30. Abomhara, M., Kien, G.M.: Cyber security and the internet of things: vulnerabilities, threats, intruders and attacks. J. Cyber Secur. Mobil. **4**, 65–88 (2015). https://doi.org/10.13052/jcsm2245-1439.414
31. Jain, A.K., Nandakumar, K., Ross, A.: 50 years of biometric research: accomplishments, challenges, and opportunities. Pattern Recognit. Lett. **79**, 80–105 (2016). https://doi.org/10.1016/j.patrec.2015.12.013
32. Yang, J., Wei, J., Shi, Y.: Accurate ROI localization and hierarchical hypersphere model for finger-vein recognition. Neurocomputing 1–11 (2018). https://doi.org/10.1016/j.neucom.2018.02.098
33. Wang, H., Xu, Z., Pedrycz, W.: An overview on the roles of fuzzy set techniques in big data processing: trends, challenges and opportunities. Knowl.-Based Syst. **118**, 15–30 (2017). https://doi.org/10.1016/j.knosys.2016.11.008
34. Eissa, M.M., Elmogy, M., Hashem, M.: Rough granular computing knowledge discovery models for medical classification. Egypt. Informatics J. **17**, 265–272 (2016)

35. Ma, Z., Li, W., Wang, L., Xu, H., Zhu, J.: X-ray astronomical point sources recognition using granular binary-tree SVM. In: International Conference on Signal Processing Proceedings, ICSP, pp. 1021–1026 (2017). https://doi.org/10.1109/ICSP.2016.7877984
36. Bezdek, J.C.: Pattern Recognition with Fuzzy Objective Function Algorithms. Plenum Press, New York, NY (1981). https://doi.org/10.1007/978-1-4757-0450-1
37. Ojala, T., Pietikäinen, M., Mäenpää, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans. Pattern Anal. Mach. Intell. **24**, 971–987 (2002). https://doi.org/10.1109/TPAMI.2002.1017623

Convolutional Neural Network for Satellite Image Classification

Mohammed Abbas Kadhim and Mohammed Hamzah Abed

Abstract Multimedia applications and processing is an exciting topic, and it is a key of many applications of artificial intelligent like video summarization, image retrieval or image classification. A convolutional neural networks have been successfully applied on multimedia approaches and used to create a system able to handle the classification without any human's interactions. In this paper, we produce effective methods for satellite image classification that are based on deep learning and using the convolutional neural network for features extraction by using AlexNet, VGG19, GoogLeNet and Resnet50 pretraining models. The Resnet50 model achieves a promising result than other models on three different dataset SAT4, SAT6 and UC Merced Land. The accuracy of classification of this model for UC Merced Land dataset is 98%, for SAT4 is 95.8%, and the result for SAT6 is 94.1%.

Keywords Satellite image classification · Deep learning · Convolutional neural network · Features extraction

1 Introduction

In recent years, remote sensing technologies have been developed quickly. That means, acquiring an extensive collection of remote sensing images with high resolution have become much more accessible. Based on this notion, many researchers of remote sensing recognition and classifications have been moving from traditional methods to recent techniques. The traditional methods depend on the intensity of pixel level interpretation while the modern techniques are focused in the semantic understanding of the images. The semantic understanding aims to classify the data

M. A. Kadhim (✉) · M. H. Abed
College of Computer Science and IT, Computer Sciences Department, University of Al-Qadisiyah, Diwaniyah, Iraq
e-mail: mohammed.abbas@qu.edu.iq

M. H. Abed
e-mail: mohammed.abed@qu.edu.iq

M. Huk et al. (eds.), *Intelligent Information and Database Systems: Recent Developments*, Studies in Computational Intelligence 830,
https://doi.org/10.1007/978-3-030-14132-5_13

into a set of semantic categories and a set of classes depending on remote sensing image content [1–3]. The image classification can be divided into three main classes according to its features [1]. The ‘handcrafted feature-based method’ focuses on different properties such as colors and shape information, which are possible properties of sense images [2–4], while ‘unsupervised feature learning-based methods’ aim to learn a set of basic functions such as a bag of words model that is used for features encoding. The most common encoding method is called quantization, and more effective method is fisher encoding, where the input in the Fisher method is a set of handcrafted characteristics, and the output is a set of learned features [5–7]. Finally, the ‘deep feature learning-based methods’ which is called Deep Learning (DL) [8–10]. In recent years, deep learning of remote sensing image features has shown an impressive capability for classification by selection of appropriate features for the problem of remote sensing image classification [1]. Selection of appropriate the deep learning is a subfield of machine learning based on multiple layers of learning. The deep learning structure extends from the classic Neural Network (NN) by adding more layers to the hidden layer part. There are many architectures of deep learning, one of them is a Convolutional Neural Network (CNN). The CNN is widespread and has been used in recent years for handling a variety and complex problems such as image recognition and classification by using a sequence of feed-forward layers. The CNN is similar to the traditional neural network, and it is made by neurons that have learnable weights and biases. The neurons receive a set of inputs and performing some non-linear processing, and it can be considered as a feed-forward artificial neural network [11]. Convolutional network architectures use the images as inputs which allow the encoding of certain properties into the architecture. The typical structure of CNN is a series of layers including a convolutional layer, a pooling layer, and full connection layers [1]. It can be said that it is a special case of the neural network that consists of one or more convolutional layers that are responsible for extracting low-level features such as lines, edges and corners; pooling/subsampling layers that make the features robust against distortion and noise; non-linear layers that work as a trigger function to signal different identification of likely features on each hidden layer; and fully connected layers that mathematically sum up a weighting of the previous layer of features [12]. In this paper, four effective strategies and architecture of CNN have been proposed to improve the performance of satellite images classification, four approaches of CNN (AlexNet, VGG19, GoogLeNet and Resnet50) have been used as a pre-trained for features extraction, each of them trained on imageNet dataset. We evaluate our methods by combining the earlier features with more in-depth features in a fully connected layer and compare all the results of the models with several novel methodologies on three datasets SAT 4, SAT6 and UCMD.

The structure of this paper is organized as the following: in Sect. 2, we present the related works with CNN for image classification and recognition. Section 3 gives an overview of the datasets that used in our system. The proposed work and its components have been discussed in Sect. 4. Finally, Sects. 5 and 6 contain the experiment results and conclusions of this work respectively.

2 Related Works

Classification of the satellite image is a process of categorizing the images depend on the object or the semantic meaning of the images so that classification can be categorized into three major parts: methods that are based on low features, or the other methods that are based on high scene features [13]. The first method of classification that are depend on low features is used a simple type of texture features or shape features, the most common methods of low features is local binary pattern or features based on histogram same as with paper [14], the researcher in that paper used the texture with LBP as a classification tool. The methods based on mid features are suitable for a complex type of images and structure [5]. The methods that are based on high features compare with other can be considered the most effective methods for complex images. The CNN is one of the most and widely used in deep learning algorithm with image processing [9].

Saikat Basu, Sangram Ganguly, and others proposed method that is a learning framework for satellite imagery "DeepSat", they focus on classification based on deep unsupervised learning "Deep Belief Network for classification" with Convolutional Neural Networks and achieve accuracy result 97.946 for SAT4-dataset and 93.916 for SAT6-dataset [10]. Ju et al. [15] produce a research paper for investigated of a widely used ensemble approaches for image classification and recognition tasks using deep convolutional neural networks. These approaches include majority voting, the Bayes Optimal Classifier, and super learner. Albert et al. [16] analyze patterns in land use in urban neighborhoods by using large-scale satellite imagery data and state-of-the art computer vision techniques basing on deep CNN. They obtain ground truth land by using class labels carefully sampled from open-source surveys, in particular, the Urban Atlas land classification dataset of 20 land use classes across 300 European cities. They also show that the deep representations extracted from satellite imagery of urban environments can be used to compare neighborhoods across several cities. Robinson et al. [17] proposed a deep learning convolutional neural networks model for creating high-resolution population estimations from satellite imagery. The proposed CNN model has been trained to predict population in the USA at a 0.01×0.01 resolution grid from 1-year composite Landsat imagery. The CNN model evaluated and validated in two ways: quantitative and qualitative. In quantitative validation, the proposed model's grid cell estimates aggregated at a county-level comparing with several US Census county-level population projections, and qualitatively, by directly interpreting the model's predictions in terms of the satellite image inputs. In general, the proposed model is an example of how machine learning techniques can be a useful tool for extracting information from inherently unstructured, remotely sensed data to provide practical solutions to social problems. Pratt et al. [18] propose Convolutional Neural Networks approach for Diabetic Retinopathy (DR) diagnosis from digital fundus images and classify its severity. They develop CNN architecture and data augmentation which can identify the intricate features that involved in the classification task such as micro-aneurysms, exudate and hemorrhages on the retina and consequently provide a diagnosis automatically without user input. They trained the proposed CNN approach using a high-end graphics processor unit (GPU) on the Kaggle dataset and demonstrate exciting results. In this work, we will focus on CNN as a classification method. Shamsolmoali et al. [19] proposed have a new classification pipeline to facilitate a high dimensional multimedia data analysis basing on a unified deep CNN and the modified residual network which can be integrate with the other feed-forward network style in an endwise training fashion.

3 Datasets

In the proposed work, we will use three different dataset SAT4, SAT 6 and UCMD. The first two types "SAT4 and SAT6" images are extracted from the NAIP program, this data set consists of 330,000 scenes spanning of all United States images. The images consist of 4 layers red, green, blue and Near Infrared (NIR). The third dataset is UC Merced Land Use Dataset contain "tif" file image format.

3.1 SAT 4

This version of the dataset consists of 500,000 image patches that are covering four lands included barren land, trees, grassland and a class that are contain all land cover classes. 400,000 classes are chosen for the training set, and the 100,000 remain are used for a testing dataset. All images are normalized into 28×28 pixels [10].

3.2 SAT6

This version of the dataset contains 405,000 images each of size 28×28 pixels, and covering six land classes barren land, trees, grassland, roads, buildings and water bodies. 324,000 images are choosing as a training dataset, and the remain 81,000 are used for testing dataset [10]. Also, Fig. 1 shows samples image of SAT 4 and SAT 6 datasets.

3.3 UC Merced Land

This dataset consists of 21 classes land use image dataset each class contains 100 image each image measures 256×256 pixel. The images extracted manually from large dataset images from the USGS National Map Urban Area Imagery collection. Figure 2 shows selected samples of the images from 20 class [20].

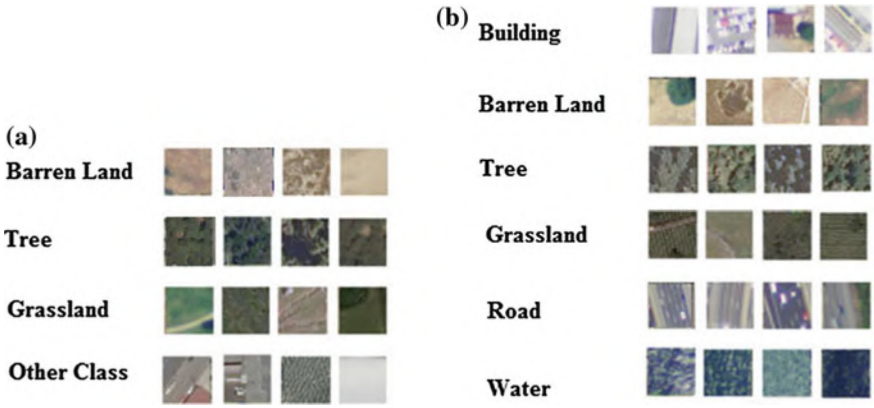


Fig. 1 Sample images “28 × 28 × 4” from **a** SAT4 and **b** SAT6 dataset

Fig. 2 Sample images from UC Merced Land dataset

4 Proposed Work

The structure of the proposed work was planned after studying the literature work on satellite image classification as in Fig. 3 that illustrates a general overview of the proposed model of satellite image classification that based on CNN. The proposed work is divided into two parts: the training phase and testing phase. The datasets are divided into two sets initially the first one is used as a training image and the second one used for testing of our models. SAT dataset consists of SAT4 and SAT6 each one contains 400,000, 324,000 images are selected as a training set consecutively and 100,000, 81,000 images are selected as a testing set.

The other datasets UC Merced Land Use that contain 21 class each one has 100 images, we have selected 70 images as training set and 30 images as testing set for all the classes. Moreover, because of the model implemented and tested on two different

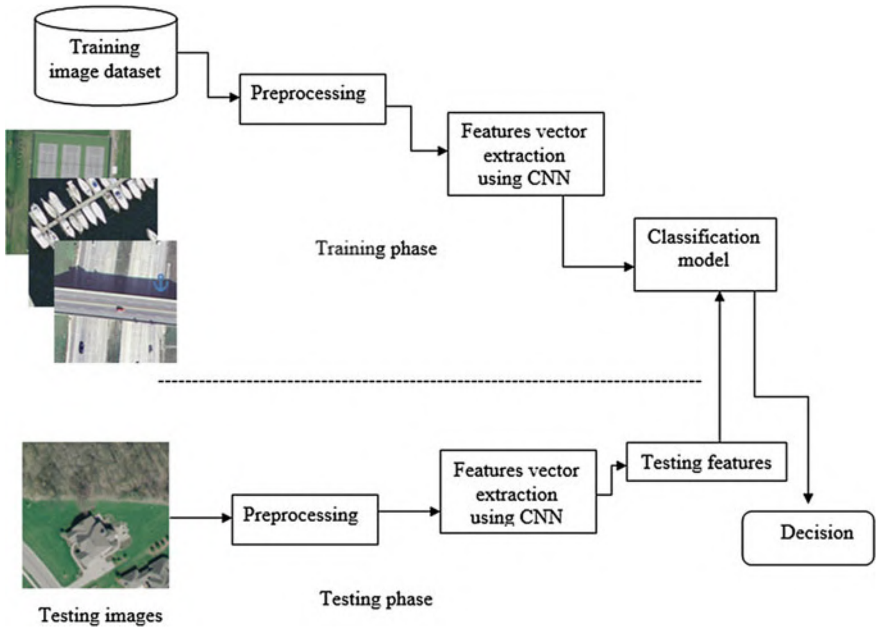


Fig. 3 The diagram of the proposed system

datasets, the preprocessing phase is such an important step to make the input images sharing the same characteristics.

4.1 Training Phase

The first stage in our model is the training phase. In this part, the selected images from both of datasets as training images are going through steps starting from preprocessing features vector extraction based on CNN.

Preprocessing. The datasets that used in our model are different, the color images of SAT airborne datasets consist of four bands 28 × 28 uint8, and the other dataset UCMD 256*256 uint8 three bands for red, green and blue. So to build a model that used for classification you must be starting with color normalization all image by reducing the invisible band NIR of the SAT datasets, convert the images into grayscale, and then the entire satellite images are ready to the next step for extracting features vector that belongs to each image in training set.

Features extraction based on CNN. The efficiency of satellite image classification is based on the power of the features that extracted from the training dataset. The power of that features will be reflected on testing phase. So by proposed off-the-

Table 1 Pretrained network, layers and features layers

	AlexNet	VGGNet-19	GoogleNet	Resnet50
Input data	Input image 227 × 227 × 3	Input image 224 × 224 × 3	Input image 224 × 224 × 3	Input image 224 × 224 × 3
Layers	25 × 1 nnet.cnn.layer.Layer	47 × 1 nnet.cnn.lay er.Layer	144 × 1 nnet.cnn.layer.Layer	177 × 1 nnet.cnn.layer.Layer
Features layer	fc8	fc8	loss3-classifier	fc1000

shelf features extraction from the images, we provide high-level features to be set of training data to train the CNN.

In this work, we have used several pretrained networks all of them have been trained on the ImageNet dataset as can visit the link http://www.image-net.org/ which have one thousand object categories. Table 1 shows the characteristic of each one that used and the fully connected layer that we have considered it as a features vector. Every CNN's layer produces an activation or response to an input image. In all these layers there are only a few layers within CNN architecture that can be suitable for features extraction of the input image. The feature that have been extracted from the deeper layer can be used as a training feature because it gives advance features contrariwise the beginning layer of the CNN capture only the primary image features like edge and blobs. The first layer of the CNN has learned for detecting the edge and blob features, and these original features are processed by deeper layer in this case the first features are combined with more in-depth high-level features in full connections layer, that can be used in recognition or classification tasks, so the fully connected layer is chosen to be features's layer.

4.2 Testing Phase

The second phase of the satellite image classification model is a testing phase. In this part, the 30% remaining of each dataset will be tested to check and measure the accuracy of the classifier method. Same as with a prepare the input data for training phase it will occur the testing images starting with preprocessing and extract set of features for all categories in the datasets and save it as two-dimensional matrices each row belongs to the one image. we will explain the experimental result of the satellite image classification based on CNN.

5 Experimental Results

We evaluate the performance of the satellite image classification that is used datasets which mention in the datasets section above. In this part, we will discuss the experimental results that are implemented based on a combination of deep features and earlier features of CNN by using four models AlexNet, VGGNet-19, GoogleNet and Resnet50 which are pretrained on imageNet dataset. The features are extracted from different layer based on the model type and full connection layers have shown in Table 1. Due to we have used different datasets and varying dimensions, we kept the size of an image and normalized the four bands into visible layers only red, green and blue. The features layer are selected in four models from last pooling full connection layer: AlexNet is layer number 23 “fc8”, VGGNet-19 is 45 layer “fc8”, googleNet is layer number 142 “loss3-classifier” and Resnet50 is layer number 175 “fc1000”. Table 2 shows the configuration of the four models on UCMD dataset.

In this work, we have tested four pretrained CNN with their configuration that are listed in Table 2, on different datasets SAT 4, SAT 6 and UC Merced Land. Each dataset is divided randomly into two part: training and a testing subset of images, Table 3 shows the datasets setting in our experimental results.

The proposed method that is based on combination of deep features and earlier features with Resnet50 that extracted from “fc1000” layer achieve better result than features extracted from first convolutional or deep convolutional, also it shows better performance than other pretrained convolutional neural network like AlexNet, VGG-19 and GoogleNet because the feature that extracted from Resnet50 are deeper than the others under the selected percentage 70% of training with the configuration of Resnet50 that shows in Table 2.

Also, Fig. 4 shows the accuracy and achieve of the Resnet50 model has the better result than other models, and Fig. 5 shows the loss of training of the samemodel in 250 epochs both of them by using UC Merced Land Datasets. Figure 6 presents the comparison among the models that used for features extraction, its visible that the Resnet50 model used for features extraction has a better result of classification than other models and loss function is less than others.

So Table 4 show the accuracy of all datasets that used with different models and algorithms.

As shown above in Table 4, the accuracy values that produced by the research paper [10] is achieved a classification ratio on SAT4 and SAT6 reached to 97.946 and 93.916 respectively. They suggested a mechanism for extracting data and features of an input image and used the principle of normalization of that features as a vector in Deep Belief Network for classification. They presented two datasets SAT4 and SAT6 and that proposed work didn’t test on UC Merced Land. The second research paper in Table 4 [21] that investigated in our experiments, the researchers proposed an agile CNN architecture named SatCNN for HSR-RS image scene classification. Based on recent improvements to modern CNN architectures and they are used a smaller kernel with effective convolutional layers to build an effective and fast CNN. The method tested on SAT4 and SAT6 with achievement ratio 99.65 and 99.54

Table 2 Configuration of the pretrained models

AlexNet		VGGNet-19		GoogleNet		Resnet50	
Layer	Configuration	Layer	Configuration	Layer	Configuration	Layer	Configuration
Conv1	Filter 96 11 × 11 × 3 Stride 1 [4 4] Pooling 3 × 3 Stride2 [2 2]	conv1_1	Filter 64 3 × 3 × 3 Stride1 [1 1] Pooling 2 × 2 Stride2 [2 2]	conv1-7 × 7_s2	Filter 64 7 × 7x3 Stride1 [2 2] Pooling 3 × 3 Stride2 [2 2]	conv1	Filter 64 7 × 7 × 3 Stride 1 [2 2] Pooling 3 × 3 Stride2 [2 2]
Conv2	Filter 256 5 × 5 × 48 Stride 1 [1 1] Pooling 3 × 3 Stride2 [2 2]	conv2_1	Filter 128 3 × 3 × 64 Stride1 [1 1] Pooling 2 × 2 Stride2 [2 2]	conv2-3 × 3_reduce	Filter 64 1 × 1 × 64 Stride 1 [1 1] Pooling 3 × 3 Stride2 [2 2]	res2a_branch2a	Filter 64 1 × 1 × 64 Stride 1 [2 2] Pooling 3 × 3 Stride2 [2 2]
Conv3	Filter 384 3 × 3 × 256 Stride 1 [1 1] Pooling 3 × 3 Stride2 [1 1]	conv3_1	Filter 256 3 × 3 × 128 Stride1 [1 1] Pooling 2 × 2 Stride2 [2 2]	inception_3a-1 × 1	Filter 64 1 × 1 × 192 Stride 1 [1 1] Pooling 3 × 3 Stride2 [1 1]	res3d_branch2c	Filter 512 1 × 1 × 128 stride [1 1]

(continued)

Table 2 (continued)

AlexNet		VGGNet-19		GoogleNet		Resnet50	
Layer	Configuration	Layer	Configuration	Layer	Configuration	Layer	Configuration
Conv4	Filter 384 $3 \times 3 \times 192$ Stride 1 [1 1] Pooling 3×3 Stride2 [1 1]	conv4_1	Filter 512 $3 \times 3 \times 256$ Stride1 [1 1] Pooling 2×2 Stride2 [2 2]	inception_3b-1 × 1	Filter 128 $1 \times 1 \times 256$ Stride1 [1 1] Pooling 3×3 Stride2 [1 1]	res4e_branch2c	Filter 1024 $1 \times 1 \times 256$ Stride1 [1 1]
Full connection 1	fc7 4096 fully connected layer	Full connection 1	fc7 4096 fully connected layer	–	–	–	–
Full connection 2	fc8 1000 fully connected layer	Full connection 2	fc8 1000 fully connected layer	Full connection 1	loss3-classifier 1000 fully connected layer	Full connection 1	fc1000 1000 fully connected layer

Table 3 Dataset setting for experimental results

Dataset	Subset of images		Images resolution	
	Training	Testing	Width	Height
SAT 4	400,000	100,000	28	28
SAT 6	324,000	81,000	28	28
UC Merced Land	21 class × 70	21 class × 30	256	256

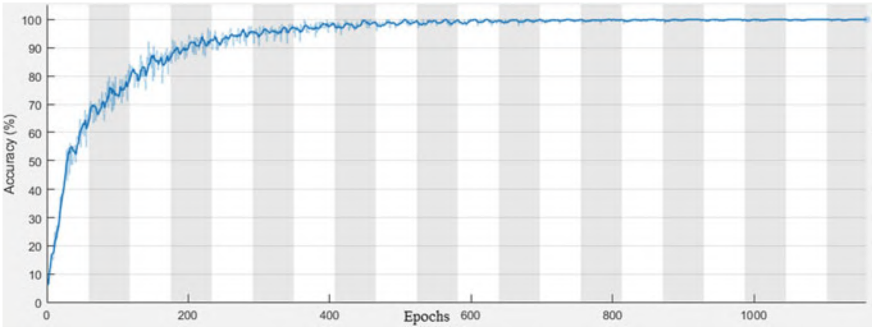


Fig. 4 Training accuracy of the Resnet50 model

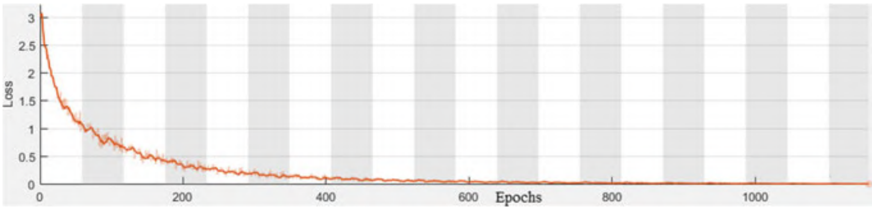


Fig. 5 The training loss of Resnet50 model

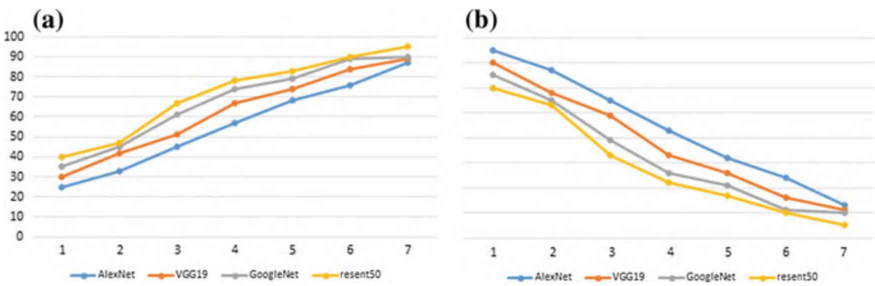


Fig. 6 Comparison between models based on UC Merced Land dataset. **a** Accuracy of training. **b** Loss of training

Table 4 Accuracy of the SAT4, SAT6 and UC Merced Land dataset based on different models and algorithms that are used

Algorithms	Classifier accuracy % SAT4	Classifier accuracy % SAT6	Classifier accuracy % UC Merced Land
DeepSat [10]	97.946	93.916	–
Agile CNN SatCNN [21]	99.65	99.54	–
Triplet networks [22]	99.72	99.65	97.99
Features extraction based on AlexNet	84	82	87
Features extraction based on VGG19	89	84	89
Features extraction based on GoogleNet	91	89	90
Features extraction based on Resnet50 (proposed method)	95.8	94.1	98

respectively and it is not tested on UC Merced Land. The performance accuracy of the third research paper [22] is 99.72, 99.65, and 97.99 on SAT4, SAT6, and UC Merced Land datasets respectively. The researchers produce a novel classification method via triple networks. In our experiment results on proposed methods based on features extraction depend on Resnet50 achievement produce the best model for classifying image set of UC Merced Land dataset. The performance of our proposed model (Resent50) is better than results yielded from research paper [10] for SAT6 dataset and it is the worst for SAT4 dataset.

6 Conclusions

In this paper, we present useful models for satellite image classification that are based on convolutional neural network, the features that are used to classify the image extracted by using four pretrained CNN models: AlexNet, VGG19, GoogleNet and Resnet50 and compare the result among them. The features are extracted from a combination layer or full connection layer of earlier layers and deep layers. After the experiment result of the datasets and the pretrained models, the Resnet50 model achieves a better result than other models for all the datasets that are used “SAT4, SAT6 and UC Merced Land”. The result of the classification based on Resnet50 as features extraction has better accuracy and minimum loss value than other methods and able to work on different datasets. The achievement of our proposed method based on Resnet50 is better result than research paper [10] for image classifying of SAT6, it is also a better than research paper [22] for classify UC Merced Land dataset.

References

1. Cheng, G., Li, Z., Yao X., Guo, L., Wei, V.: Remote sensing image scene classification using bag of convolutional features. IEEE Geosci. Remote Sensing Lett. **14**(10), (2017)
2. Bian, X., Chen, C., Tian, L., Du, Q.: Fusing local and global features for high-resolution scene classification. IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens. **10**(6), 2889–2901 (2017)
3. Huang, L., Chen, C., Li, W., Du, Q.: Remote sensing image scene classification using multi-scale completed local binary patterns and Fisher vectors. Remote Sens. **8**(10), (2016)
4. Chen, C., Zhang, B., Su, H., Li, W., Wang, L.: Land-use scene classification using multi-scale completed local binary patterns. Signal Image Video Process. **10**(4), 745–752 (2016)
5. Zhang, F., Du, B., Zhang, L.: Saliency-guided unsupervised feature learning for scene classification. IEEE Trans. Geosci. Remote Sens. **53**(4), 2175–2184 (2015)
6. Li, Y., Tao, C., Tan, Y., Shang, K., Tian, J.: Unsupervised multilayer feature learning for satellite image scene classification. IEEE Geosci. Remote Sens. Lett. **13**(2), 157–161 (2016)
7. Yuan, Y., Wan, J., Wang, Q.: Congested scene classification via efficient unsupervised feature learning and density estimation. Pattern Recogn. **56**, 159–169 (2016)
8. Yao, X., Han, J., Cheng, G., Qian, X., Guo, L.: Semantic annotation of high-resolution satellite images via weakly supervised learning. IEEE Trans. Geosci. Remote Sens. **54**(6), 3660–3671 (2016)
9. Zou, Q., Ni, L., Zhang, T., Wang, Q.: Deep learning based feature selection for remote sensing scene classification. IEEE Geosci. Remote Sens. Lett. **12**(11), 2321–2325 (2015)
10. Basu, Saikat, Ganguly, Sangram, Mukhopadhyay, Supratik, DiBiano, Robert, Karki, Manohar, Nemani, Ramakrishna: DeepSat—A Learning Framework For Satellite Imagery, SIGSPATIAL'15, Nov 03–06, 2015. Bellevue, WA, USA (2015)
11. Yu, X., Wu, X., Luo, C., Ren, P.: Deep learning in remote sensing scene classification: a data augmentation enhanced convolutional neural network framework. GISci. Remote Sensing (2017)
12. Hijazi, S., Kumar, R., Rowen, C.: Using Convolutional Neural Networks for Image Recognition, IP Group, Cadence
13. Yu, Y., Liu, F.: Dense connectivity based two-stream deep feature fusion framework for aerial scene classification. www.mdpi.com/journal/remotesensing (2018)
14. Ojala, T., Pietikainen, M., Maenpaa, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans. Pattern Anal. Mach. Intell. **24**, 971–987 (2002)
15. Ju, C., Bibaut, A., van der Laan, M.J.: The relative performance of ensemble methods with deep convolutional neural networks for image classification, ArXiv e-prints, Apr (2017)
16. Albert, A., Kaur, J., Gonzalez, M.: Using convolutional networks and satellite imagery to identify patterns in urban environments at a large scale. In: Proceeding of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining pp. 1357–1366 (2017)
17. Robinson, C., Hohman, F., Dilkina, B.: A deep learning approach for population estimation from satellite imagery. In: Proceedings of the 1st ACM SIGSPATIAL Workshop on Geospatial Humanities, pp. 47–54 (2017)
18. Pratt, H., Coenen, F., Broadbent, D.M., Harding, S.P., Zheng, Y.: Convolutional neural networks for diabetic retinopathy. In: International Conference On Medical Imaging Understanding and Analysis, MIUA 2016, Loughborough, UK, (2016)
19. Shamsolmoali, P., Jain, DK., Zareapoor, M., Yan, J., Alam, M.A.: High-dimensional multimedia classification using deep CNN and extended residual units. Multimedia Tools Appl. https://doi.org/10.1007/s11042-018-6146-7 (2018)
20. Yang, Y., Newsam, S.: Bag-of-visual-words and spatial extensions for land-use classification. In: ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (ACM GIS), (2010)

21. Zhong, Yanfei, Fei, Feng, Liu, Yanfei, Zhao, Bei, Jiao, Hongzan, Zhang, Liangpei: SatCNN: satellite image dataset classification using agile convolutional neural networks. Remote Sensing Lett. **8**(2), 136–145 (2017)
22. Liu, Yishu, Huang, Chao: Scene Classification via Triplet Networks. IEEE J Selected Topics Appl Earth Observ. Remote Sensing **11**(1), 220–237 (2018)

A Survey of Tacit Knowledge in Community Learning Based on Measurement Dimensions

Winanti, Ford L. Gaol, Raymond Kosala, Suhono Harso Supangkat and Benny Ranti

Abstract Tacit Knowledge is one of the potential assets in developing community-based learning. A problem often encountered is the difficulty in formalizing and articulating the knowledge, and for this reason, a media is required to extract into an easily translated form. This study focuses on explicit forms of knowledge, and as a result, the method employed is the collection of research findings related to tacit knowledge and community learning. Also, these findings were surveyed based on predetermined keywords and analysed into a complete report which can be used as material for decision making especially in the matter of sharing this knowledge in community-based learning. The purpose of this study was to determine the understanding of measuring this knowledge in terms of existing dimensions, such as technical and cognitive. However, in this case, the conditions and processes help inform about tacit with reference to the Community. The survey activity is still limited to its conceptual measurement and has not yet led to implementation in the form of community learning where members can interact directly using text, images, videos

Winanti (✉)
BINUS Graduate Program—Doctor of Computer Science, Bina Nusantara University, West Jakarta 11480, Indonesia
e-mail: winanti12@ipem.ac.id

Winanti
Information System Department, STMIK Insan Pembangunan, Tangerang, Indonesia

F. L. Gaol · R. Kosala · B. Ranti
Computer Science Department, Doctor of Computer Science, Bina Nusantara University, West Jakarta 11480, Indonesia
e-mail: fgaol@binus.edu

R. Kosala
e-mail: rkosala@gmail.com

B. Ranti
e-mail: benny.ranti@gmail.com

S. H. Supangkat
System and Information Technology Department, Institute Technology Bandung, Bandung, Indonesia
e-mail: suhono@stei.itb.ac.id

M. Huk et al. (eds.), *Intelligent Information and Database Systems: Recent Developments*, Studies in Computational Intelligence 830,
https://doi.org/10.1007/978-3-030-14132-5_14

and other forms. In addition to using two dimensions of measurement, knowledge can be shared through socialization, externalization, combination and internalization that refer to the Nonaka and Ba models for learning to provide a sense of fun for the community group.

Keywords Survey · Tacit knowledge · Community learning

1 Introduction

Knowledge management is an activity carried out by a person in transforming tacit into explicit knowledge asking for knowledge to be created from individual sources based on experience, reasoning, intuition, one's personal judgment and insight. The former knowledge is a very strategic asset and has the potential to compete [1].

The world of education is a collection of knowledge through daily activities, as it manages to collect, store, share, produce and utilize in a structured manner in accordance with the existing curriculum. However, it has not been able to benefit from tacit knowledge [2].

Acknowledging the significance of tacit knowledge in everyday life in creating new sources of knowledge that were initially unthinkable by humans to become a context that can be utilized and truly felt. In the last few years, the community has focused on knowledge management without realizing that this knowledge in question has not evolved and does not make sense. But over time, it has become a new knowledge that makes sense and is very beneficial to human life [3].

This Knowledge for a group of communities is one that is unwittingly able to provide new knowledge about things that have not been seen or structured yet. This knowledge is important and useful for some community groups that have no idea about schooling, because it provides inspiration and new knowledge.

In various communities where members communicate with one another, the average information they convey to the discussion forum is from the experience and habits of their day-to-day life. Humans cannot be totally independent. They need to interact with others, even if just an interaction among neighbours. However, it is the nature of a relationship and social network one part that often becomes a significant aspect of a community. Emerging communities make use of information technology innovations which consist of a set of organizations having the purpose of adopting, implementing and assimilating the use of information technology and networks connected online [4].

This paper aims to determine the application of tacit knowledge in the context of community-based learning, and therefore, the obtained result is expected to provide recommendations useful for the wider community to utilize this knowledge with the intention of disseminating community-based knowledge. This paper is still limited to the conceptual survey of the measurement of tacit knowledge seen from two dimensions of measurement (technical and cognitive dimension), although there are several other dimensions that can be used. The implementation of the dimensions will be carried out by the authors in the next study and will not be discussed in this paper.

2 Literature Review

Tacit knowledge is a type obtained from personal knowledge and understanding, and cannot always be articulated in code. It is very useful in sharing experiences and knowledge which have originated from actions, routines or habits, strong commitment, ideas, and values of science [3].

The community itself can be interpreted as a collection of members who have a sense of belonging and bound among one another with the belief that their needs will be fulfilled as long as they are committed to continue together [5].

The gathering of actors in the community dynamically evolves, as a form of mutual attention to developing innovations and activities that occur in society is a "general project". Members of the innovation community come from various ethnicities, tribes and religions, but are all grouped into the same interests, having the same motivation and goals [4, 5].

In the 21st century, humans are faced with life in a knowledge based community and society, where knowledge serves as a source of strength for humans who have high quality of life and survival. As a result, the effects of market, products, technology, competition and regulation change very rapidly. Continuous innovation due to rapid technological development is carried out by testing existing theories and knowledge in the form of spirals to pass contradictory concepts such as the concept of tacit and explicit knowledge.

Community-based learning components consist of (a) community members i.e. components that are directly involved in community learning, and can either be an individual or a group. (b) Collaborative learning, usually emphasizing on initiative not being a result of engineering and thinking of other people to cooperate. Cooperation and collaboration have almost the same meaning, namely cooperation, which distinguishes only cooperation with direction and assistance from others or cooperation carried out on its own initiative. (c) Learning problems, namely a learning approach that always begins with the presentation of problems designed in a context that is relevant to the material being studied (problems adapted to the material). (d) Learning practices, namely learning efforts carried out by giving participants the opportunity to have firsthand experiences. The community process comprises of (a) identifying problems, (b) exploring alternatives, (c) working together in taking action, and (d) revising and refining the plan [4, 6].

Chergui et al. (2018) stated that knowledge includes hidden elements that cannot be reduced and are related to the personal nature of individuals, and that they are very difficult to formalize, communicate and share. Remembering tacit knowledge usually comes from personal actions or attitudes [7].

Nonaka and Takeuchi (1995) stated that the application of knowledge management is the process of creating knowledge that arises from the interaction of its conversion process called SECI (socialization, externalization, combination and internalization) [3] (Fig. 1).

According to Nonaka, the conversion of knowledge can be performed with the following statement:

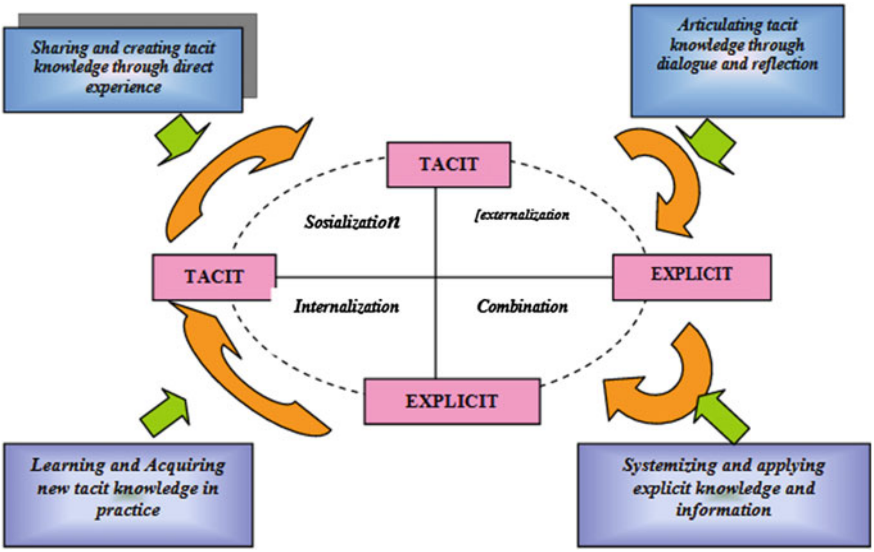


Fig. 1 Knowledge conversion process SECI. *Source* Nonaka et al. [3]

a. Socialization is the conversion of tacit-tacit knowledge to individuals and groups. This is the process of disseminating experience, and the creation of knowledge through spreading mental models and technical skills.
b. Externalization is tacit-explicit conversion at the organizational level between organizations and groups. This is the core of the knowledge creation process in form of a hypothesis concept or model.
c. Combination is an explicit conversion that occurs at the group level—individual, or among individuals. This is a process of systemizing concepts in a knowledge system.
d. Internalization is an explicit conversion—tacit. It is the process of converting explicit knowledge to tacit knowledge.

According to Polanyi (1966), tacit knowledge has the following properties: [8]

a. Cannot be shared.
b. They are things that are more widely known than delivered.
c. It often consists of habits and cultures that cannot be determined by themselves.
d. Cannot be codified, but can be transferred or obtained from experience.
e. Describe the know what (facts) and know why (science).
f. Engage learning and skills.
g. Formed in organizational groups and relationships, core values, assumptions and beliefs that are difficult to identify, store, count or map.

The Nonaka model provides a few guidelines for specific actions which can be taken to facilitate the creation of knowledge by converting, linking and combining them into a different knowledge in order for the required activities to be adapted into generative learning [3, 9].

Knowledge is a combination of information and experience [10, 11]. The average unstructured understanding, experience and knowledge has not been well documented, considering that knowledge is still within the limits of one's expertise and experience. For this reason, it is necessary to set up and explain therefore that their understanding, experience and expertise can be translated into a structured language that will be easy to understand, learn and practice to support the already existing knowledge [10].

Knowledge is grouped into two, and they are tacit and explicit knowledge. There are interactions between both groups that help create new ideas which are difficult. Despite the fact that there are researches which classify knowledge as individual or group background, practical or theoretical, hard or soft, internal or external, the background tacit and explicit classification remains the most common used by many researchers which is also very practical [3, 12].

3 Methodology

The method is a series of ways, processes or steps used in asserting a scientific field. Understanding of Survey was the method used in conducting a detailed and complete literature review. There are three steps taken, namely: the first stage defines the scope in detail, and determines the research question; the second stage is conducting a search that involves selecting a literature source based on keywords; and the final stage is by determining the inclusion and exclusion criteria to sort reference types for further exploration [13].

The purpose of the Survey is firstly to assist the author in understanding the problems being examined in accordance with the frame of mind; secondly, to find out the theory, findings and other material to be used as a basis for the research; thirdly, to form a framework for solving a problem.

The author gathers references as literature, using the following steps: Firstly, the author looks for literature with keywords to answer questions on "Survey Tacit Knowledge in Community Learning Based on Measurement Dimensions". The second step is to focus on publications in work papers, the results of conferences, as well as proceedings and journals published from 2000 to 2018. The third step is the literature writer who focuses on tacit knowledge and community learning.

The scope of this research is sharing tacit knowledge implementation at community learning. The first step taken by the author is to put together a review which defines literature sources in order to obtain articles correlating with the topic, sourced from journals, proceedings and research results. The literature review source is systematically based on: (a) Science Direct (www.sciencedirect.com), (b) Link Springer (link.springer.com), (c) DOAJ or Directory Open Access Journal (http://doaj.org), (d) Google Scholar (scholar.google.com), (e). IEEEXplore Digital Library (http://ieeexplore.ieee.org). Useful keywords in the search of this paper are: ("sharing"

OR "Contribution") AND "Tacit Knowledge" AND ("Dimension" OR "Part") AND ("Community" OR "Association") AND ("Learning" OR "Discuss").

From several references found, the author can put together a summary and come to a conclusion that basically, this knowledge can be measured through different dimensions, and these dimensions can be used in determining the type that will be used in the community-based learning process, given that this learning cannot be separated from tacit knowledge. Furthermore, community members can exchange ideas to solve problems they face. Strong ties make them encourage one another in matters of improvement and problem solving. However, it is often difficult to translate the knowledge delivered to a structured form because their ideas and views prioritise the spontaneity they feel at that time, based on their guesses/prejudices and personal understanding. The following is sharing tacit knowledge with measurements made in reference to the two measurement dimensions seen in Fig. 2.

Following Nonaka (2000) and the book Sangkala (2007), the authors can make measurements of tacit knowledge with models that are easy to understand and implement. Two dimensions are needed in sharing this knowledge based on community learning, and they are:

1. Technical dimensions: These dimensions are in form of skills that are difficult to formalize, and they are usually owned by people proficient in a particular field, but there are limitations in the educational sector.

   a. Tacit knowledge can be measured by subjective judgments of individuals. For some people, subjective attitudes must be upheld to view a problem where a person thinks relatively. Furthermore, the result of guessing, based on people's feelings or tastes, as limited as estimates and assumptions are supported by facts/data.
   b. Personal understanding is an individual opinion which cannot be codified, but only be obtained from experience.
   c. Intuitive thinking is a cognitive process that generates ideas as a strategy in making the right decisions expected to produce spontaneous problem solving answers.
   d. Suspicion or prejudice is to make a decision before knowing the relevant facts about the object. Prejudice itself has three types, namely cognitive prejudice which refers to what is considered right; affective prejudice refers to what is liked and disliked, co native prejudice refers to how a person tends to act.
   e. Inspiration is a process that encourages or stimulates the mind to carry out a task or activity, especially something creative. It is often referred to as the mind that has been attached to the soul or heart.
   f. Idea is a result of thoughts arranged in the human mind. It has an unlimited scope, and each individual has a diverse idea that can be replicated in the form of new knowledge to complement each other's existing knowledge.

2. Cognitive dimensions, those that are difficult to articulate consist of:

   a. Trust attitudes shown by humans when they feel enough to know and conclude that they have reached the truth.

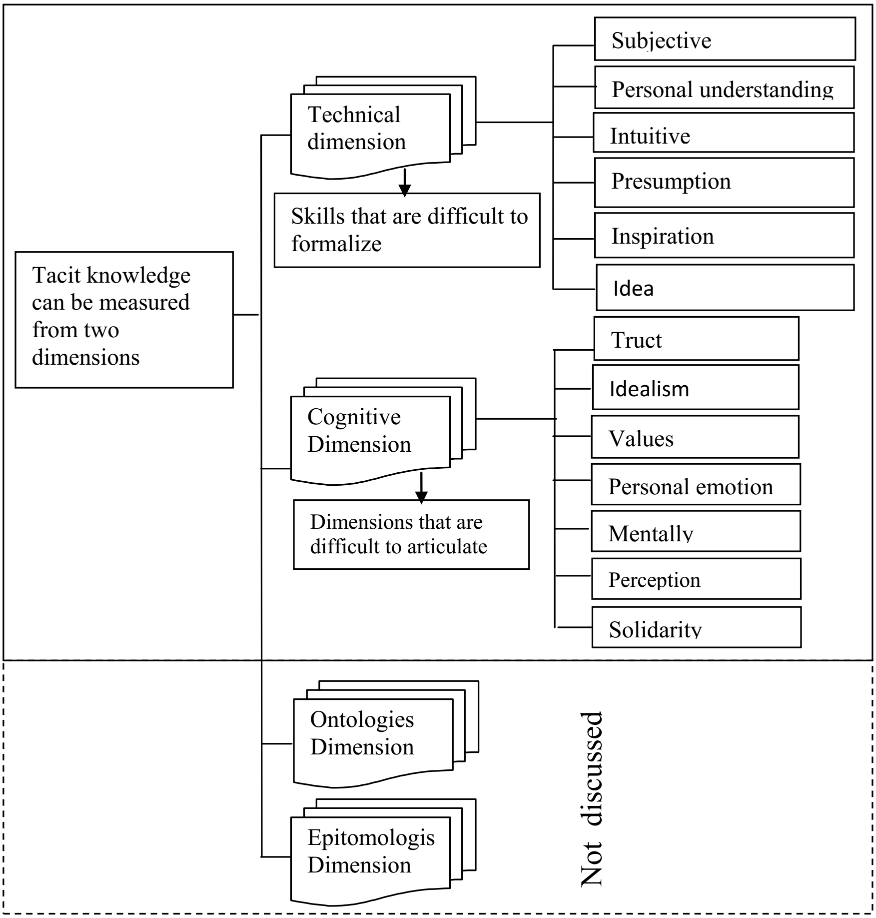


Fig. 2 Tacit knowledge sharing with dimensions based on community learning

b. Idealism is a belief in matters that are considered to be right based on experience, education, cultural culture and habits. It grows slowly in one's soul, and manifests in forms of behaviour, attitudes, ideas or ways of thinking.
c. Value contains elements of consideration which bring an individual's ideas about things that are right, good, or desirable.
d. Emotions are shown when feeling happy about something, angry with someone, or afraid of something.
e. Mental is related to thoughts, intelligence, memories or processes that have to do with thoughts, reasons, and memories [14].
f. Perception is "an individual process in interpreting, regulating and giving meaning to stimulus that comes from the environment in which s/he is located, and it is the result of the learning and experience process" [15].

g. Solidarity is an attitude owned by humans as regards the expression of human feelings for the sense of the same fate and affinity for other people and groups in their community.

Based on the explanations given above, the writer adds one component, called solidarity. However, it has not been discussed in the technical and cognitive dimensions according to the two authors mentioned above. This component is very important in tacit measurement, because it is very synonymous with cooperation among individuals in a particular group. The sense of belonging and affinity for others in their group cannot be expressed but felt.

4 Result and Discussion

E-learning systems can improve knowledge management for company activities in order to provide benefits such as models being presented in such a way as the system integration and knowledge management. Both models are applied to several real-world scenarios to describe benefit integration, and participants can become virtual communities [16].

Developing an understanding of the conditions and process of sharing tacit knowledge in the context of community-based learning using naturalistic questions, and in the process, methodology finds that online learning encourages processes and creates conditions that are consistent with the knowledge creation models [3, 17]. Transfer knowledge through checking tacit knowledge and providing insights that make it quiet. The type of tacit knowledge is identified through: Skill, Cause-effect, Composite, Cognitive, Cultural, Unlearning, Human, Taboo, and Emotional. The various mechanisms used in transferring this knowledge experienced by each unique set of elements become a tacit [18].

Managing this knowledge effectively and efficiently is the key successful organizational measurement factors by paying attention to several measurement dimensions to gain competitive advantage. This is also carried out with the use of the knowledge which must be well understood because it is not only difficult to understand, but also one of the most challenging in knowledge management. It is a valuable strategic resource for a community and organization with sustainable potential and competitive advantage in the future. For a better understanding, the level of tacit knowledge is divided into four factors, namely: factors from individual/personal, managerial, expertise, and collective knowledge factors. The influence of managerial knowledge impacts the greatest influence on tacit knowledge among all four factors, and then verbal communication is the most important item in this factor. Collective and task knowledge have the same effect [12].

Community learning is very important for individuals and communities because it encourages collaboration, emergence of new skills and life skills, willingness to live with natural resources effectively with awareness of the individual in accordance with what they do. This process begins with finding problems experienced by community members. After a member identifies a problem, they start exchanging opinions, conducting discussions, analysing the cause of the problem, finding solutions, and furthermore taking appropriate steps and actions to face the problems [6].

Participating in the learning community virtually has many advantages and benefits such as access to information, performance achieved by increased creativity and professional identity to be developed for the better. But in this case, the integration is rich in a comprehensive knowledge management model at the institutional level [7, 19, 20]. Supported by virtual learning networks in the context of community learning in building capacity for the purpose of developing partnerships between national and international institutions to obtain better and more advanced knowledge in the future [20].

Knowledge in the context of tacit and certain relationships between consciences that cannot be manifested together and usually made through concrete actions in terms of the relationship between the two becomes very important [7].

Conducted by training, social interaction with knowledge sharing was carried out among community members. The tutor does not have to be interacted with directly, since the sharing can also be done via online media. Community members can only communicate with an android cellphone, laptop or personal computer connected to the internet. The dimensions used in the measurement of tacit knowledge in community learning are seen from the technical dimensions and cognitive dimensions, as shown in Fig. 2. Actions are taken towards this measurement:

(1) Knowledge exists when someone has existing coupled with his personal knowledge and is developed in accordance with the advancement of science and technology, which is then shared with community members using text, images, or videos.
(2) Identification of actors involved in community learning. The determinants and benchmarks of its success are the involvement and activeness of the actors.
(3) After identifying the actors involved, set the goals and then choose the top priority to predict satisfaction of community members.
(4) Every member of the community can exchange responses, answers and objections to the questions and problems that are conveyed to the forum, so that they can all convey ideas based on their knowledge and experience to perfect their answers.
(5) Satisfaction of community-based learning can finally be achieved with a good view of the members who exchange answers and responses for tacit to be seen from the technical dimension which was originally difficult to formalize into interest and complete. Measurement of tacit knowledge seen from a cognitive point of view is easier because of the many and complete input from community members.

The measurement of tacit knowledge was carried out by survey [21] and online media was used in data collection. It was collected in an extractive manner using software in order to generate a completely new knowledge from the community members. This method becomes very interesting for community members in addition to solving a problem. Many jobs that require organizational knowledge, intellectual capital, communities or organizations that create new knowledge stem from an understanding of its nature [22]. It is sometimes difficult to find, but often found when people experience a problem or event. The features are often found in the real world, but cannot be displayed [23]. Due to this fact, it is necessary to measure tacit knowledge while still looking at the technical and cognitive dimensions so that it can be translated into complete, useful and fun information for community groups.

5 Conclusion and Implication

The emergence of various communities spread throughout the world gets rid of distance as an obstacle for those who want to exchange ideas. The constraints they face today are in terms of sharing knowledge which is still carried out using unstructured methods. Methods and models are therefore needed to transform or extract from unstructured to structured knowledge, or better known as tacit (unstructured) to explicit (structured) knowledge. This paper prioritizing tacit knowledge sharing is seen from a measurement perspective by using the technical and cognitive dimension in the perspective of the knowledge for community-based learning which is currently growing.

Based on the results of the survey conducted using several previously conducted studies and papers, were measured from two main dimensions. The first is technical dimension that is subjective, personal, intuitive, guess/prejudice, inspiration and ideas. The second is the cognitive dimension that originates from beliefs, idealism, values, one's emotions and mental state, and perceptions. Also, authors added solidarity in the cognitive dimension where solidarity is very important in tacit measurement especially in terms of cooperation between individuals and groups that have been difficultly articulated. These two dimensions often occur and are experienced in sharing tacit knowledge in community-based learning. Problems experienced and observed from community members can be organized into groups so that other members can provide input, opinions or ideas to resolve these problems. Strength and cohesiveness among community members make them big and an inseparable part of tacit knowledge with community learning supported by various social media that they can use to share problems and share ideas quickly and easily.

6 Limitation and Future Research

This paper discusses knowledge in the context of community learning based on two dimensions. The first is the technical dimension which is in a complicated form, difficult to understand and formalize. The second is the cognitive dimension which can be interpreted as one that is difficult to articulate in the form of structured knowledge. The authors argue that these two dimensions are very important in developing community-based learning. Others such as ontologism and etymological dimensions are not discussed in this paper. The author also did not discuss the extraction of tacit knowledge by means of socialization, externalization and combination. Other dimensions will be discussed in subsequent studies, and the extraction from unstructured into structured forms of knowledge was also left out in this paper.

Two dimensions that have not been discussed will be in the next study, either by the authors themselves, or by other writers who will conduct researches on the same topic, so that a complete research and tacit knowledge can be measured from 4 dimensions intact namely: technical, cognitive, ontological and the epistemological dimensions. With this, learning in a community can be carried out in a compelling and interesting way.

References

1. Ozmen, F.: The capabilities of the educational organizations in making use of tacit knowledge. **9**, 1860–1865, (2010)
2. Fullan, M.: The role of leadership in the promotion of knowledge management in schools. OECD Conference (18–19 March 2002). Retrieved on 6 May 2010
3. Nonaka, I., Toyama, R., Konno, N.: SECI, Ba and Leadership : a Uni® ed model of dynamic knowledge creation. **33**, 5–34 (2000)
4. Wang, P., Ramiller, N.C.: Community learning in IT innovation. MIS Q. **33**(4), 709–734 (2009)
5. McMillan, D.W., Chavis, D.M.: Sense of community: a definition and theory. Am. J. Commun. Psychol. **14**(1), 6–23 (1986)
6. Techavijitsarn, W., Kovitaya, M., Ratana-ubol, A.: The development of a community learning model for self-management in conserving community forests. Proc. Soc. Behav. Sci. **191**, 2325–2328 (2015)
7. Chergui, W., Zidat, S., Marir, F.: An approach to the acquisition of tacit knowledge based on an ontological model. J. King Saud Univ. Comput. Inf. Sci. (2018)
8. Polanyi, M.: The Tacit Dimension. Routledge and Kegan Paul, London (1967)
9. Arling, P.A., Chun, M.W.S.: Facilitating new knowledge creation and obtaining KM maturity. **15**(2), 231–250 (2011)
10. Alavi, M., Leidner, D.E.: Management review: knowledge systems: management knowledge and foundations conceptual. Manag. Inf. Syst. **25**(1), 107–136 (2001)
11. Chen, Y.: Development of a method for ontology-based empirical knowledge representation and reasoning. Decis. Support Syst. **50**(1), 1–20 (2010)
12. Zaim, H., Faruk, Ö., Tar, M., Zaim, S., Alpkan, L.: Determining the critical factors of tacit knowledge in service industry in Turkey. **207**, 759–767 (2015)
13. Meyliana, Hidayanto, A.N., Budiardjo, E.K.: The critical success factors for customer relationship management implementation: a Systematic literature review. Int. J. Bus. Inf. Syst. **23**(2), 131–174 (2016)
14. Chaplin, C.P.: Kamus Psikologi, terj, Kartini Kartono, Jakarta: PT grafindo Persada, p. 407 (1995)
15. Ali, M., Asrori, M.: Psikologi Remaja Pengembangan Peserta Didik. Edisi 6. PT. Bumi Aksara, Jakarta (2009)
16. W. Q. Qwaider: Integrated of knowledge management and E- learning system **4**(4), 59–70 (2011)
17. Tee, M.Y., Karney, D.: Sharing and cultivating tacit knowledge in an online learning environment. pp. 385–413 (2010)
18. Dinur, A.: Tacit knowledge taxonomy and transfer : case-based research. Feb 2015

19. Strunga, A.: The integration of virtual learning communities into universities' knowledge management models. Proc. Soc. Behav. Sci. **197**, 2430–2434 (2015)
20. Boulay, R., van Raalte, L.: Impacting the science community through teacher development: utilizing virtual learning. Int. J. Technol. Knowledge Soc. **9**(4), 13–24 (2014)
21. Cheng, L.V., Helen, H., Jin, Z.: The effect of parental opportunism, IJV â€™ s autonomy and tacit knowledge on IJV instability: a comparison of multi-variate regression and fuzzy-set qualitative comparative analysis. J. Bus. Res. **69**(11), 5203–5209 (2016)
22. Cook, S.D.N., Brown, J.S., Cook, S.D.N., Seely, J.: Bridging Epistemologies : the generative dance between organizational knowledge and organizational knowing. May 2015, 1999
23. Isoni, A., Poulsen, A., Sugden, R., Tsutsui, K.: Focal points in tacit bargaining problems: experimental evidence. Eur. Econ. Rev. **59**, 167–188 (2013)

Land Cover Classification Based on Sentinel-2 Satellite Imagery Using Convolutional Neural Network Model: A Case Study in Semarang Area, Indonesia

Yaya Heryadi and Eka Miranda

Abstract Regional land use planning and monitoring remain an issue in many developing countries. Efficient solution for both tasks depended on remote sensing technology to capture and analyze remotely sensed data of the region of interest. Although a plethora of methods for land cover classification have been reported, the problem remained a challenging task in computer vision field. The advent of deep learning method in the past decade has been very instrumental to develop a robust method for land cover classification using satellite imagery as input. The objective of this paper was to present empiric results on using CNN as a land cover classifier model using Sentinel-2 spatial satellite imagery. Prior to model training, the input image representation was extracted using eCognition to produce texture, brightness, shape, and vegetation index. Land cover labeling followed the Land Cover Class in Medium Resolution Optical Imagery Interpretation document provided by Indonesian National Standardization Agency. The training of CNN model achieved 0.98 mean training accuracy and 0.98 mean testing accuracy. As comparison, the same data and same feature were trained with another model: Gradient Boosting Model (GBM). The results revealed that the training accuracy and testing accuracy with GBMs were 0.98 and 0.95 respectively. CNN model showed small improvement of the accuracy to classify land cover with the image feature (NDVI, Brightness, GLCM homogeneity and Rectangular fit).

Keywords CNN · Land cover classification · Satellite imagery

Y. Heryadi (✉)
Computer Science Department, BINUS Graduate Program—Doctor of Computer Science, Bina Nusantara University, Jakarta, Indonesia
e-mail: yayaheryadi@binus.edu

E. Miranda
Information Systems Department, School of Information Systems, Bina Nusantara University, Jakarta, Indonesia
e-mail: ekamiranda@binus.ac.id

M. Huk et al. (eds.), *Intelligent Information and Database Systems: Recent Developments*, Studies in Computational Intelligence 830,
https://doi.org/10.1007/978-3-030-14132-5_15

1 Introduction

Regional land use planning and monitoring remain an issue in many developing countries such as Indonesia. In the past two decades, both tasks largely depended on remote sensing technology to capture and analyze remotely sensed data of the region of interest. Remote sensing is a technology to obtain and analyze information over the earth. This technology facilitates faster data acquisition and wider data coverage compared with conventional technology or field observations [1]. The development of remote sensing technology was followed by the development of digital image analysis. Wide access to satellite imagery in the past ten years has attracted wide interest of researchers from various research domains into remote sensing imagery analytics, such as land cover classification, due to its wide potential applications [1].

Land use classification is a prominent computer vision problem that has gained wide research interest for many years. The challenge in addressing this task is how to duplicate human abilities in image understanding, thus the computer could recognize an object like a human. The advent of deep learning [2] in the past decade coupled with wide availability of satellite imagery and GPU computer server have regained research interest on satellite image analysis. These studies resulted in a plethora of proposed models available in literature. Many of the previous studies have reported a great result of deep learning model for satellite image analysis. One of the models that have been widely adopted is convolutional neural network (CNN) which is designed to imitate the image recognition system on human visual cortex [2]. To achieve such objective, the CNN model was designed as a hierarchy of convolutional and pooling layers and a fully connected layer (FCL) [3]. The hierarchical structure of CNN makes it possible for the model to learn hierarchical feature of the input data. In remote sensing image analysis, another critical factor for CNN performance is satellite image features.

The feature of satellite imagery can be divided broadly into two feature categories that represent: low-level objects and high-level objects. In the past decade, many research works mainly focused on classifying individual pixels or objects by identifying low-level local image features [4], such as color histogram [5]. Although having been widely used for satellite imagery classification, low-level local image features only represent object's local features such as color, shape, and texture. Consequently, the classifier can merely classify low-level objects, e.g. road, and soil [6]. On the other hand, land cover imagery requires to recognize multiple and more complex objects such as settlement, forest, water body and bare land. For that purpose, the feature of land cover imagery, therefore, should represent a wide range of objects. This requirement brought about the need of a method to extract various features automatically from a satellite image. The study by [2] claimed the strength of deep learning model was its ability to learn hierarchical feature from input data. Convolutional neural network (CNN) model is one of the deep learning models which is capable to perform such ability.

Despite many studies have proposed CNN model for land cover classification, to the best of our knowledge, little have been said on the use of spectral satellite image as input. Therefore, the objective of this study is to build a CNN model as land cover classifier using Sentinel-2 satellite imagery as input.

2 Literature Review

2.1 Convolutional Neural Network Model

Convolutional Neural Network (CNN) is a term referring to a neural network model which has been specifically used to process grid-structured data i.e. two-dimensional image. Convolutional Neural Network (CNN) is a development of multilayer perceptron (MLP) designed to process two-dimensional data in the image form [1]. Following [2], as a deep learning model, CNN architecture can be built with various deep structure. Despite its structure depth, the mapping process of input to output in a CNN model follows several stages. Each stage produces a data representation called feature maps. The structure wise of each stage consists of three layers, namely convolution, layer activation function, and pooling layer [1]. The Convolutional Neural Network architecture network is showed in Fig. 1.

In general, the first stage in the CNN architecture is the convolution stage uses a kernel of a certain size. The number of kernels used depends on the number of features produced. The second stage is the activation function; this function usually uses the ReLU activation function (Linear Unit Rectifier). Subsequently after passing the activation function, the process goes through the pooling process. This process is repeated several times until a sufficient feature map is obtained to proceed to the fully connected neural network [2].

The effective way to classify high resolution images by combining deep features and image objects has gained wide research interest. For example, a study reported by Zhao et al. [7] proposed a method utilizing deep CNN framework to automatically extract robust and discriminative features for the classification of complex urban

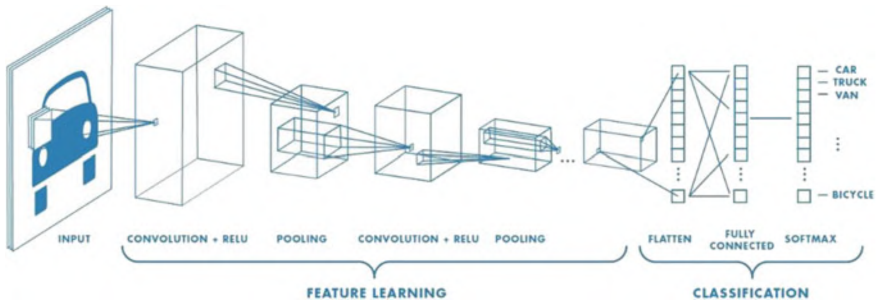


Fig. 1 A typical CNN achitecture. *Source* [2]

objects (such as building roofs and cars). However, the study has no access to the contextual information at the global level.

Kussul et al. [8] presented the use of CNN to reach the target accuracy of 85% for major crops (wheat, maize, sunflower, soybeans, and sugar beet). This study became a foundation for further operational use of remote sensing data within Sentinel-2 imagery.

Zhu and Newsam [9] demonstrated that land use classification is possible using high-level image features extracted using CNNs from geo-located ground-level images. The framework showing that high-level, semantic image features extracted using pre-trained CNN models generalize well to related problems.

Yang et al. [10] investigated the use of different encoder-decoder structures of CNN based on SegNet for the pixel-wise classification of land cover based on aerial images and derived data. The experiments have shown that an ensemble of CNN having different architectures and using different input data achieves the best performance with an overall accuracy of almost 86% for eight land cover classes (building (build.), sealed area (seal.), bare soil (soil), grass, tree, water, car, etc.). Kampffmeyer and Jenssen [11] proposed revision of a convolutional neural network architecture for urban remote sensing image segmentation trained on data modalities which are not all available at test time.

Suzuki et al. [12] proposed a Convolutional Neural Network (CNN) which mimics professional interpreters' manual techniques. With the proposed CNN, K. Suzuki shows that the multi-modal CNN works robustly and gets more than 80% user's accuracy. A novel Deep Learning architecture to leverage PAN and MS imagery for land cover classification has been proposed by [13]. The proposed method, MultiResoLCC, consists of a two-branch end-to-end network which extracts features from each source at their native resolution and lately combines them to perform land cover classification at the higher-resolution single-band panchromatic PAN resolution.

2.2 Convolutional Layer

Convolution layer is a layer in CNN architecture that performs a convolution operation on the output of the previous layer. This layer is the main process underlying the CNN architecture network. Convolution operations are operations on two functions of real value arguments. This operation applies the output function as a feature map from the input image. This input and output can be seen as two value arguments [2]. Convolutional operation can be seen in Eq. 1.

$$s(t) = (x * w)(t) = \sum_{a=-\infty}^{\infty} x(a)w(t-a). \quad (1)$$

where: s(t) = function as a result of convolution operation; x = input; w = weight (kernel). The function s(t) provides a single output in the form of a map feature. The first argument is the input (x) and the second argument (w) is the kernel or filter. When input as a two-dimensional image, it can be said t as a pixel and replace it with i and j.

3 Research Method

3.1 Study Area and Data

The test site for this study is Semarang area, Central Java, Indonesia. This area is selected for the following reasons: (1) Cloud covers of the observed area (2) Characteristics of geography and topography of Semarang area represent the land cover class analyzed in this study (3) It is one of the cities in Indonesia where the population growth and the development of the city has caused the increase of land requirement. Such reasons make regional land use planning and monitoring are crucial.

The Sentinel-2 satellite imagery was retrieved through web site with URL https://earthexplorer.usgs.gov/. The data was acquired on August 27, 2017, at 14:55:47 with the cloud cover in the observed area at 8.49%, Sun Zenith Angle Mean (average sun exposure in the observation area in vertical angle) = 29.93 and Sun Azimuth Angle Mean (average sun exposure in the observation area in horizontal angle) = 55.98. The image acquired was not in the rainy season, with high exposure of sun, dry and very low humidity.

The main land cover classes for this study were based on the Land Cover Classification of SNI 7645: 1: 2014 National Standardization Agency of Indonesia and National Standardization Agency of Indonesia for Land Cover Class in Medium Resolution Optical Imagery Interpretation consisting of: (1) primary dry forest, (2) secondary dry forest, (3) planting forest, (4) grassland, (5) settlement, (6) water body, and (7) bare land. These land cover classes were used in this study.

The Sentinel-2 is used for the following reasons: (1) This satellite image has 13 bands obtained from the MSI (Multispectral Imager) instrument, (2) Temporal resolution of Sentinel-2 is 10 days performed by one satellite and 5 days performed with two satellites that will make large amounts of observational data produced. This satellite has spatial resolution from 10 to 60 m. The satellite image is composed of the following bands: 4 (red), bands 3 (green), bands 2 (blue), bands 8 (Near-Infrared) and bands 11 (SWIR, Short-Wave Infrared). The band 4 is useful for identifying types of vegetation, soil and urban features; band 3 provides excellent contrast between clear and turbid (muddy) water; band 2 is useful for land and vegetation identification, forest type mapping, and to identify human-made features; while band 11 is useful for measuring soil moisture and vegetation, and it provides good contrast between

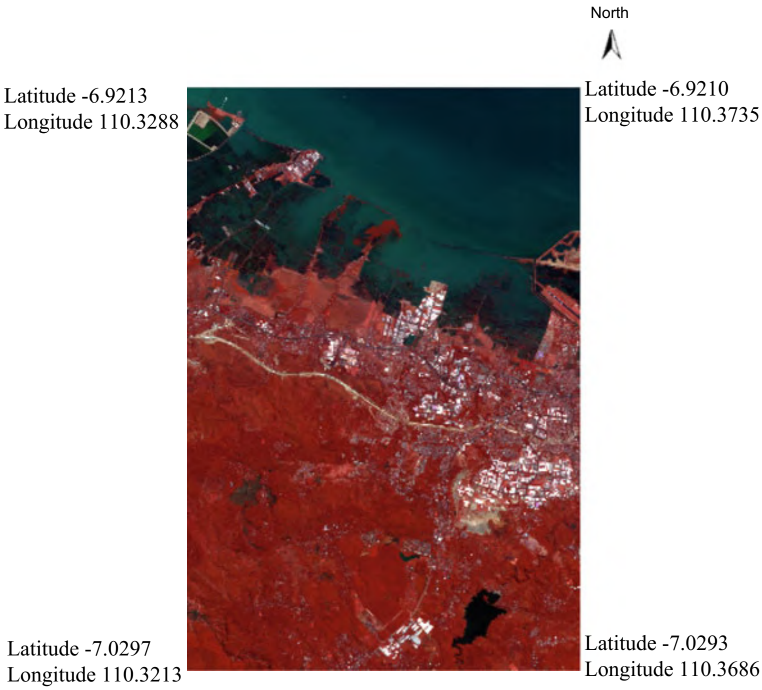


Fig. 2 False color image fusion result (band 4, 3, 2) of the Sentinel-2 satellite imagery for Semarang, Central Java, Indonesia

various types of vegetation. Figure 2 shows false color image fusion result (band 4, 3, 2) of the Sentinel-2 satellite for Semarang, Central Java Province, Indonesia.

3.2 Land Cover Classes

The main land cover classes for this study were based on the Land Cover Classification of SNI 7645: 1: 2014 National Standardization Agency of Indonesia and National Standardization Agency of Indonesia for Land Cover Class in Medium Resolution Optical Imagery Interpretation. The Land Cover Classification of SNI 7645: 1: 2014 National Standardization Agency of Indonesia was compiled based on the UNFAO (Food and Agriculture Organization of the United Nations) land cover classification system and ISO 19144-1—Geographic Information—Classification Systems—Part 1: Classification System Structure ISO 19144-1. Following [14], definition of each land cover classes and its image features in this study can be summarized in Table 1 as follows.

Table 1 Land covers definition of each land cover class

Land cover class	Definition	Image features
Primary dry forest	The primary dry forest is characterized by the presence of dark green objects (in bands 8, 4, 3), tends to dark and coarse texture with clustered tree canopy. There are no logged marks	1. Color (band 8, 4, 3) 2. Hue 3. Texture 4. Shape
Secondary dry forest	The primary dry forest is characterized by the presence of dark green objects (in bands 8, 4, 3), tends to dark and coarse texture with clustered tree canopy. There are logged marks	1. Color (band 8, 4, 3) 2. Hue 3. Texture 4. Shape
Planting forest	Green (on the bands 8, 4, 3). Neatly arranged and have a certain pattern	1. Color (band 8, 4, 3) 2. Hue 3. Texture 4. Shape
Grassland	Characterized by thin lines of very fine textured vegetation in moss green (on the bands 8, 4, 3)	1. Color (band 8, 4, 3) 2. Hue 3. Texture 4. Shape
Settlement	Characterized by a group of dense building patterns in urban settlements and sparse building pattern in rural settlement. The road network looks solid	1. Color (band 8, 4, 3) 2. Hue 3. Texture 4. Shape
Water body	Objects are indicated by the existence of light blue area, whitish blue or black (on bands 8, 4, 3) covering a fairly wide area	1. Color (band 8, 4, 3) 2. Hue 3. Texture 4. Shape
Bare land	Objects (in bands 8, 4, 3) are characterized by pink to dark red area sometimes brown, depending on the content of the soil material, and white when the material is composed by lime	1. Color (band 8, 4, 3) 2. Hue 3. Texture 4. Shape

Source [14]

3.3 *Feature Extraction*

In this study, feature extraction of satellite image was implemented using eCognition software in several steps (see Fig. 3). First, eCognition implemented polygon-based segmentation to divide image into a number of segments/objects based on pixel similarity. This process was implemented by means of fuzzy logic that allowed the integration of a broad spectrum of different object features, such as spectral values, shape, and texture [15]. In this study, multi-resolution segmentation technique was applied to sequentially combine pixels with similar values into an object.

Next, image features were extracted from each segment to represent the associated land cover class. The extracted features are namely: Color, Hue, Texture, and Shape. Finally, each of the extracted features were transformed using eCognition to produce another set of features namely: NDVI (Normalized Difference Vegetation Index), Brightness, GLCM (the Gray-Level Co-occurrence Matrix) homogeneity, and Rectangular fit. Among the extracted features, NDVI (Normalized Difference Vegetation Index) is widely used as vegetation index due to its simplicity but effective for quantifying green vegetation. The NDVI value is calculated using the following equation [16].

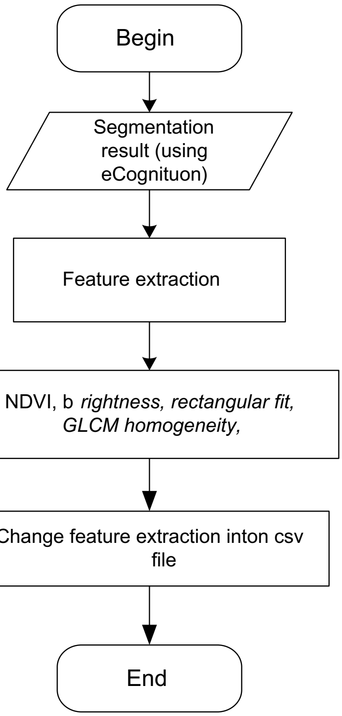


Fig. 3 Feature extraction process

$$NDVI = \frac{(B08 - B04)}{(B08 + B04)} \tag{2}$$

3.4 The Classification Process with CNN Model

The process of land cover classification can be represented by Fig. 4. In general, the process comprises of the following main steps: loading data from hard disk, data splitting into training and testing dataset, training CNN model, and evaluating model performance.

The main layer of a CNN architecture used in this study has the following formats: Conv1D (filters = 10, kernel_size = 2, activation = 'relu', input_shape = (1, 4), padding = 'same')) whose parameters and its meaning are:

1. filters = 10 is an output dimension (number of output filter in a convolutional).
2. kernel_size = 2 is the length of convolutional window.
3. activation = 'relu', activation function for Rectified Linear Unit (relu). ReLu does not activate negative input.
4. input_shape = (1, 4) is the shape of input vector.
5. padding = 'same' produces padding input, therefore output has the same length with the original input.
6. Epochs = 5000 is the number of times the entire dataset is passed through a neural network [1].

k-Fold Cross-validation (k is a constant) is a method to evaluate performance of a classifier model by first splitting the input dataset randomly into k partition. This method works in the following steps. First, the dataset is randomly split into k partitions. Second, for each unique partition, allocate the partition as a testing dataset and the remaining partitions as training dataset. The training dataset is used to fit or testing the model followed by evaluation using the testing dataset. This process is repeated k-times until each partition is used as testing dataset and the remaining

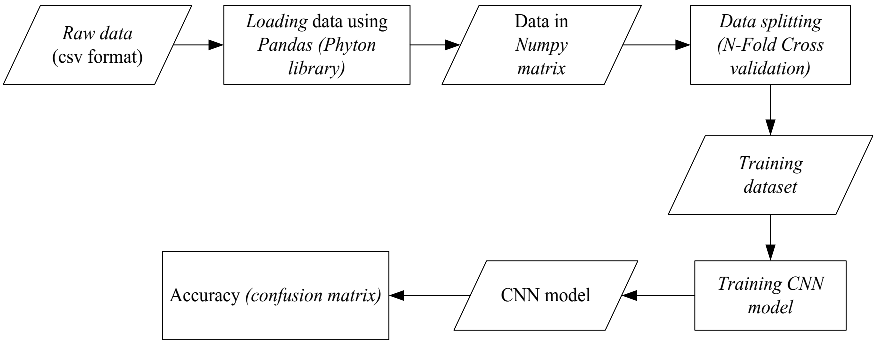


Fig. 4 Training process of the land cover classification model

partition as training dataset. Compute the average performance metrics as model performance measure [16, 17].

This study used supervised approach. Creating training samples from input image is a semi-automatic process. First, each segment/object resulted from segmentation process is represented by features (NDVI, Brightness, GLCM homogeneity and Rectangular fit). Next, compare each individual segment with the standard object in the Indonesia Thematic Map published by the Ministry of Forest and Environment 2017 for the study location to obtain a theme/class for each segment [18]. Final result of this process is a labeled dataset for training a classifier.

4 Research Result and Discussion

Image segmentation process using eCognition produced 2,072 segments/objects. For each segment as input, eCognition further produced features that represent each segment. Next, a semi-supervised process used to set out label for each segment. Finally, the overall labeled dataset was used to train classifier model supervisedly using Root Mean Square Error (RMSE) as objective function. In this study, K-Fold cross-validation is used to measure performance of two models namely: boosted tree gradient and Convolutional Neural Network (CNN) models. Model performance was measured using training, testing accuracy, producer accuracy and user accuracy.

Producer's accuracy is metric that measures accuracy of the map maker's (producer) view. The value can be interpreted as how often the real features in the field are correctly displayed on the classification map or probability of the certain land cover can be classified correctly. Producer accuracy was computed from the amount of reference data classified accurately and divided by the total number of reference data for a class. On the other hand, user accuracy is a metric that measures accuracy from the user's view, instead of mapmaker accuracy. The value of user accuracy (reliability) measures the proportion of the class on the map was actually present in the field. User accuracy was computed by taking the total number of correct classifications for a particular class divided by the total value of predictive data for a class. Table 2 shows user's and producer's accuracy of the model training.

The result of model performance evaluation of the training process was summarized in a confusion matrix (see Fig. 5). As can be seen in Fig. 5, the training accuracy is 0.984 in which it measures proportion of the training data classified by the trained model correctly.

As can be seen in Figs. 5 and 6, the training accuracy and testing accuracy of CNN model were 0.984 and 0.98, respectively. Producer accuracy showing how often the real features in the field were correctly displayed on the calcified map or probability of land cover was classified correctly. Whilst, user accuracy showed how often the class on the map will actually be present in the field (reliability). As can be seen in Tables 2 and 3, the model achieved high training and testing user accuracy and producer accuracy. It can be concluded that the trained model achieved

Table 2 User and producer accuracy of CNN training

Land cover class	Producer accuracy	User accuracy
Primary dry forest	153/157 = 0.975	153/153 = 1
Secondary dry forest	147/150 = 0.980	147/152 = 0.9671
Planting forest	147/150 = 0.980	147/150 = 0.980
Grassland	245/251 = 0.9761	245/252 = 0.972
Settlement	165/165 = 1	165/165 = 1
Water body	224/233 = 0.9614	224/230 = 0.974
Bare land	446/446 = 1	446/450 = 0.991

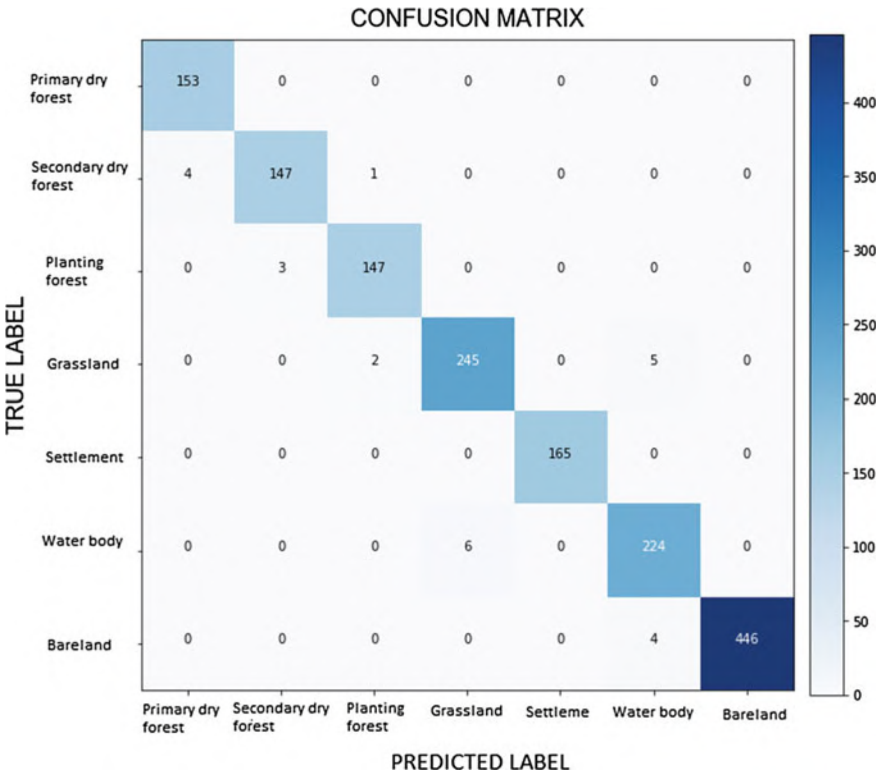


Fig. 5 Confusion matrix from CNN training

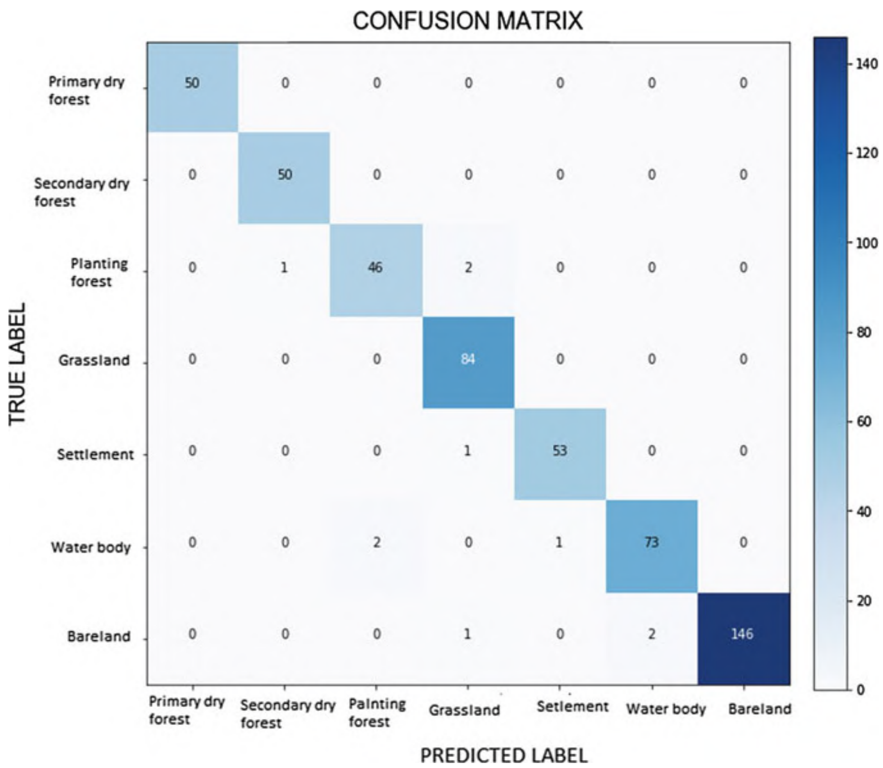


Fig. 6 Confusion matrix from CNN testing

Table 3 User's and producer's accuracy of model testing

Land cover class	Producer accuracy	User accuracy
Primary dry forest	50/50 = 1	50/50 = 1
Secondary dry forest	50/51 = 0.980	50/50 = 1
Planting forest	46/48 = 0.958	46/49 = 0.939
Grassland	84/88 = 0.955	84/84 = 1
Settlement	53/54 = 0.981	53/54 = 0.981
Water body	73/75 = 0.973	73/76 = 0.961
Bare land	146/146 = 1	146/149 = 0.980

high performance for land cover classification task or achieved a high probability to classify land cover correctly.

The results of classification reveal high accuracy when it classified land cover using CNN with the image features (NDVI, Brightness, GLCM homogeneity and Rectangular fit). Furthermore, this study examines the method employed by another artificial method.

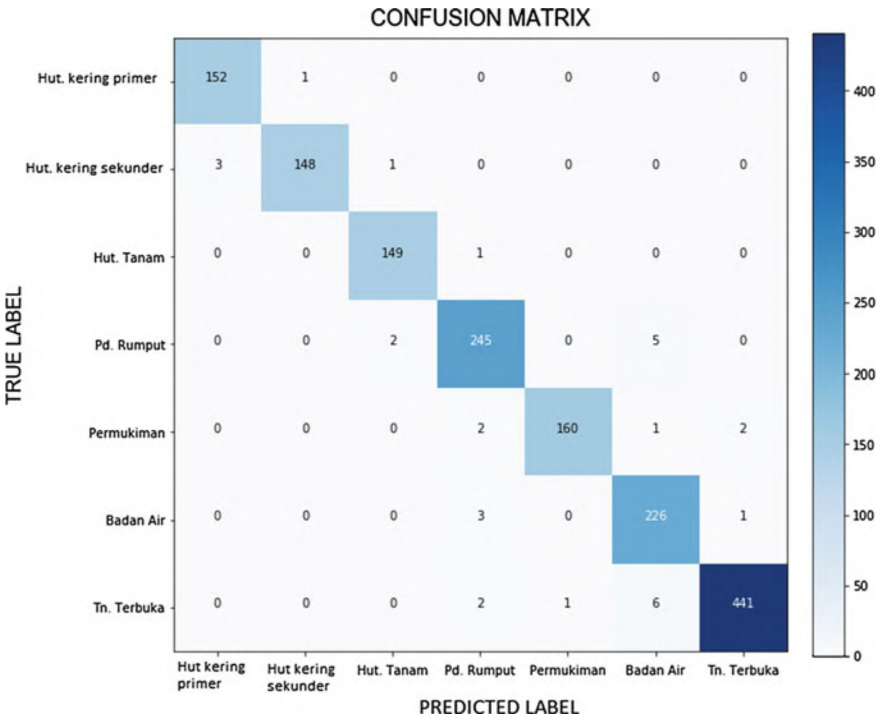


Fig. 7 Confusion matrix from GBM training

Gradient boosted tree model was used as a comparison model. The gradient boosting (GBMs) is a machine learning technique for classification problems. GBM is an extremely popular machine learning algorithm that has proven successful across many domains [19]. Figures 7 and 8 show the confusion matrix for GBM training and testing process. As can be seen in Figs. 7 and 8, the training accuracy and testing accuracy were 0.98 and 0.951, respectively.

In compare to GBM model, CNN model showed small accuracy improvement to classify land cover with the image feature: NDVI, Brightness, GLCM homogeneity and Rectangular fit.

5 Conclusion

Regional land use planning and monitoring remain an issue in many developing countries such as Indonesia. Despite many proposed models have been reported from previous studies, the task remained a challenging problem. The advent of deep learning models and wide availability of satellite imagery in the past decade has

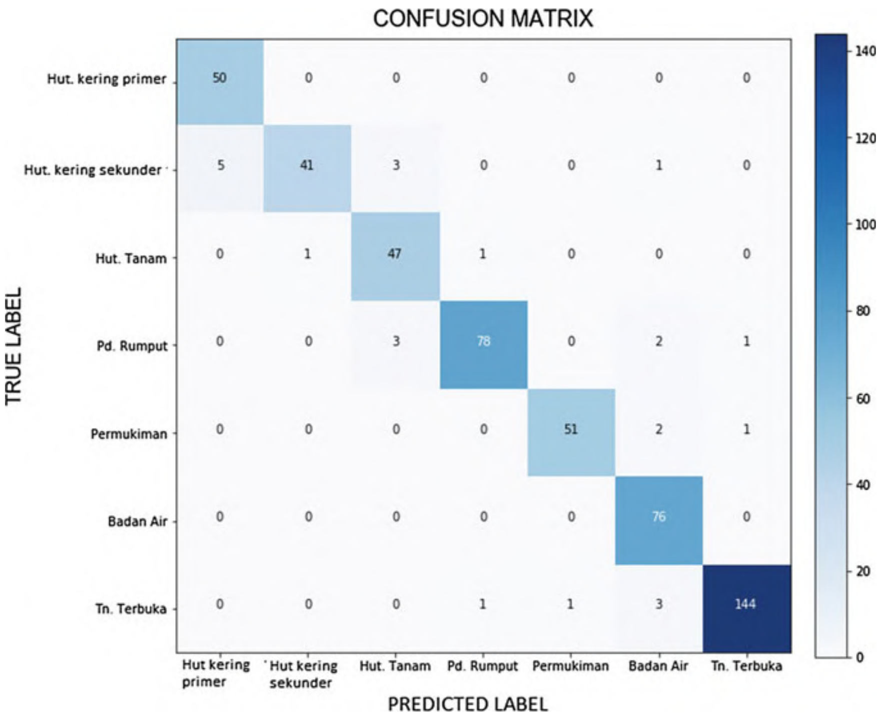


Fig. 8 Confusion matrix from GBM testing

motivated many researchers to adopt deep learning model to address land cover classification problem.

This study presented the results of land cover classification using Sentinel-2 satellite imagery as input, CNN model as the classifier, and Semarang area as the location for this study. The results showed that CNN model achieved high performance for classifying image objects dataset using 4 features namely: texture, brightness, shape, and vegetation index.

In this study, the input image was Sentinel-2 satellite imagery. eCognition was used as a tool to do the following tasks: (1) segmentation of input image into a number of polygon-based segments, (2) extract Color, Hue, Texture, and Shape as feature of each segment, and (3) transform the extracted features from each segment into NDVI, Brightness, GLCM homogeneity and Rectangular fit. A semi-supervised process was implemented to set the land cover class for each segment. Given a set of image segments represented by a set of features and a set of land cover classes associated by each segment, a deep learning model such as CNN can be trained supervisedly as land cover classifier. Hence, in principle, this method can be adopted to train land cover classification model at any location given availability of Sentinel-2 satellite imagery for the location of interest and availability of thematic map standard published by local government to determine land cover class for each segment.

The future research works will include exploration of the other deep learning models to address the problem and extracting more satellite imagery features from other area with suitable land cover classes that fit to the area of interest.

References

1. Weng, Q., Mao, Z., Lin, J., Liao, X.: Land-use scene classification based on a CNN using a constrained extreme learning machine. Int. J. Remote Sensing pp. 1–18 (2018)
2. Goodfellow, I., Bengio, Y., Courville, A.: Deep learning (adaptive computation and mechine learning series). The MIT Press (2016)
3. Dahl, J.V., Koch, K.C., Kleinhans, E., Ostwald, E., Schulz, G., Buell, U., Hanrath, P.: Convolutional networks and applications in vision. In: Proceedings of IEEE International Symposium on Circuits and Systems, pp. 253–256, Paris, France (2010)
4. Moustakidis, S., Mallinis, G., Koutsias, N., Theocharis, J.B.: SVM-based fuzzy decision trees for classification of high spatial resolution remote sensing images. IEEE Trans. Geosci. Remote Sens. **50**(1), 149–169 (2012)
5. Swain, M.J., Ballard, D.H.: Color Indexing. Int. J. Comput. Vision **7**(1), 11–32 (1991)
6. Xia, G.S., Hu, J., Hu, F., Shi, B., Bai, X., Zhong, Y., Zhang, L.: AID: a benchmark dataset for performance evaluation of aerial scene classification. IEEE Trans. Geosci. Remote Sens. **55**(7), 3965–3981 (2017)
7. Zhao, W., Du, S., Emery, W.J.: Object-based convolutional neural network for high-resolution imagery classification. IEEE J. Sel Topics Appl. Earth Observations And Remote Sensing, pp. 3386–3396 (2017)
8. Kussul, N., Lavreniuk, M., Skakun, S. Shelestov, A.: Deep learning classification of land cover and crop types using remote sensing data. IEEE Geosci. Remote Sensing Lett., pp. 778–782 (2017)
9. Zhu, Yi., Newsam, S.: Land use classification using convolutional neural networks applied to ground-level images, In: International Conference on Advances in Geographic Information Systems Proceedings on Proceedings, pp 1–4, Seattle, USA (2015)
10. Yang, C., Rottensteiner, F., Heipke, C.: Classification of land cover and land use based on convolutional neural networks. ISPRS Annals Photogr. Remote Sensing Spatial Inf. Sci. **4**(3), 251–258 (2018)
11. Kampffmeyer, M., Salberg, A., Jenssen, R.: Urban land cover classification with missing data using deep convolutional neural networks, In: IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Texas, USA (2017)
12. Suzuki, K., Rin, U., Maeda, Y., Takeda, H.: Forest cover classification using geospatial multimodal data. Remote Sensing Spatial Inf. Sci. **42**(2), 1091–1096 (2018)
13. Gaetano, R., Ienco, D., Ose, K., Cresson, R.: A two-branch CNN architecture for land cover classification of PAN and MS imagery. Remote Sensing. **10**(1746), 1–20 (2018)
14. National Standardization Agency. Indonesian national standard RSNI-1. Land cover class in medium resolution optical imagery interpretation. Indonesia
15. Feature extraction based on object oriented analysis. http://www.ecognition.com, last accessed 2018/12/03
16. Sentinel 2 EO. (n.d.). Sentinel Hub. Retrieved 10 Apr 2018, from Sentinel 2 EO products: https://sentinel-hub.com/develop/documentation/eo_products/Sentinel2EOproducts, last accessed 2018/11/28
17. Belgiau, M., Dragut, L., Strobl, J.: quantitative evaluation of variations in rule-based classifications of land cover in urban neighbourhoods using WorldView-2 imagery. ISPRS J. Photogr. Remote Sensing, pp. 205–215 (2014)

18. Indonesia Ministry of Forest and Environment (2017). WebGIS Indonesia Ministery of Forest and Environment. Retrieved 20 Apr 2018, from Peta Interaktif KLHK: http://webgis.dephut.go.id:8080/kemenhut/index.php/id/peta/peta-interaktif (2018, last accessed 2018/11/28
19. Brownlee, J.: Machine learning mastery from feature importance and feature selection with XGBoost in Python. https://machinelearningmastery.com/feature-importance-and-feature-selection-with-xgboost-in-python/. Last accesses 2018/10/15

Part III
Computer Vision Techniques and Applications

Using the Game Engine in the Animation Production Process

Artur Bąk and Marzena Wojciechowska

Abstract The latest animated films from major studios, such as Pixar and Disney, make a great impression, but not everyone is aware that the production process of such films absorbs the work of hundreds of specialists, and their budgets often exceed hundreds of millions of dollars. The way of creating these films has not changed for decades, and the quality and realism, both constantly improving as the audience needs, are achieved by constantly increasing the resources of both human and computer power. Thanks to the recent advances in hardware technology, in particular graphics cards, as well as the development of game engines oriented not only on programmers, but also on artists, ideas appeared for their use to optimize the traditional pipeline in animated film production. This paper describes the traditional production pipeline with factors causing its high costs, the concept of its optimization by game engine together with the latest practical examples and describes the most important challenges of this new trend.

1 Introduction

Both the production of advanced computer games and the production of animated films have one common goal, which are high quality synthetic 3D video sequences. Although it seems obvious that both cases can be realized in a similar way, there are significant differences that make it not so easy. The game requires real-time graphics rendering, and its quality is not always a priority, whereas an animated movie usually requires high-quality graphics that can be generated offline. Initially, in previous decades, game engines were designed to provide a flexible environment for game production by implementing a compact set of functional modules (such as Physics, Artificial Intelligence or Lighting). Unfortunately, most of the work had

A. Bąk (✉) · M. Wojciechowska
Polish-Japanese Academy of Information Technology, Warsaw, Poland
e-mail: abak@pjwstk.edu.pl

M. Huk et al. (eds.), *Intelligent Information and Database Systems: Recent Developments*, Studies in Computational Intelligence 830,
https://doi.org/10.1007/978-3-030-14132-5_16

to be still done by programming, and the low quality of the graphics in real time was unacceptable in the film domain. Gradually in the last few years, professional graphics editing tools have been introduced to game engines [1, 2], so they can also be used by artists without programming skills. Additionally, the computing efficiency of the latest hardware has been dramatically increased, enabling high-quality graphics in real time. All these factors have led to intense research and practical attempts to use game engines in the production process of animated films.

The following chapters present a brief description of the traditional animation production pipeline, the concept of the game engine as an integrated environment for game developers, and then the use of such engines for the production of films, indicating the advantages and disadvantages of such a solution. Also presented are recent cases of using game engines to create high-quality animations. Finally, there is a discussion summarizing the most interesting aspects and challenges based on conclusions reported by experts in this field.

Related work The use of game engines in the production of animated films is a relatively new field in which the first serious achievements were published no earlier than in 2016. For this reason, there have been few papers describing this new trend. One of them is work [3] describing the production process of the Fortnite teaser, in which Unreal Engine 4 [1] was used. It mainly describes the differences between using real-time and the traditional approaches observed during this process. In turn, the article [4] summarizes real-time perspectives mainly on the basis of reports published by existing creators trying this technology in films. Our work differs from both because it describes the general approach of using real-time in animated movies without focusing on a particular engine and presents conclusions regarding a new trend gathered by experts from key companies providing real-time technology and those trying to use this technology in film production.

2 Traditional Animation Production Pipeline

The standard approach for production of animated films is based on a linear process similarly as in many other manufacturing industries, where subsequent production phases are carried out sequentially. In such a process each successive phase depends on the results of the previous one and can usually be started when the previous one is completed. In very simplified terms, the whole process can be divided into the following stages: pre-production, production and post-production. A more detailed description can be found in the literature [5], while this chapter presents a short summary of these stages, their mutual dependencies and the major disadvantages and difficulties that result from such an approach.

2.1 Pre-production

The pre-production phase is a set of defined actions aimed at developing the film concept and assumptions for its editing. Mainly popular graphics tools such as Photoshop are used here. The individual results of the post-production phase include:

Shooting script—being a summary not differing from the script of the feature film. It is often possible to estimate the initial cost of the created movie.

Visual concept—based on determining how realistic the world should be created and how its elements should be matched. The concept must be adapted to the story and its mood. Among other things, it determines lighting, virtual camera behavior, movie speed, and animation style. The existing film references are often used (graphics, comics or games). In this phase, the creative process for artists is also developed, suitable for achieving satisfying results. At this stage, it is possible to estimate a more accurate delivery time and a more accurate production budget, as well as to optimize them.

Storyboard—which is a series of drawings illustrating the artistic concept of the film and can be treated as the initial information for all departments working on the animation, on the basis of which they can assess their role in production. It is also the breakdown of the scenario into the prime factors, in order to achieve the intended goal, for example a night-time scene does not require many details of the set design, while the character approximation requires more details of the characters' faces. At this stage a short story is created and the rhythm, appropriate tension, camera settings and movements are defined. There are also defined colors, lighting and a composition showing the space on which story takes place. Cuttings between scenes are also planned to establish editing assumptions.

2.2 Production

In this phase, the target characters, objects, animations and additional visual effects are modeled. In a traditional production stream animations are made for a specific camera view, and only what is visible is rendered. Modeling takes place in 3D tools such as 3ds max, Maya, Houdini or Z-Brush, while the final scene is assembled in one of them (e.g. in 3ds max). Stilomatics and animatics prepared in tools such as Adobie Premiere or Adobe AfterEffects are useful. The final scenes are rendered using tools such as V-Ray in the form of layers (separately human figures, reflections, atmospheric effects, etc.) and assembled in tools such as Nuke, Fusion or Adobe AfterEffects. The individual results of the production include:

Stilomatic—uses drawings prepared for a storyboard and is used to present a project for stakeholders and collaborators in the project.

Animatic—being a preliminary version of the film in the form of the successive storyboard sections. It determines the tempo and mechanics of production, and is a guide to what is supposed to happen and determines the final length of the film. Animatics verify the assumptions of the pre-production storyboard. Usually, there is no need to perform full animation here, but only for the necessary parts to make it possible to synchronize, position and move throughout the movie. After the approval of the animatics, the production of the final animations begins.

Soundtrack—includes preliminary dialogs to measure and adjust timing, synchronization and emotional planning. Combining them with animatics allows for verifying their quality—whether they are well-written, whether there are too many of them, or sound natural.

Modeling—creation of 3D models of characters, designing of character movements, creation of sets and scenography models, face design with effects related to its expression and lip synchronization.

Lighting—affects all elements and production aspects of animation, from giving the film a climate, and ending with costs. Determining the appropriate texture of the film allows for creating depth, atmosphere and mood.

Compositing—combining all elements of animation and effects in the final image, including depth of field, motion blur, atmospheric effects and shadows.

2.3 Post-production

Post-production is the last phase in which all elements are combined. Working soundtracks are replaced by target and final rendering begins. Ideally, the synchronizations are closed and the animatics are replaced by the final recordings. There may be corrections to remove or swap shots. In addition, a color correction is carried out, which in the simplest form is the color alignment between the shots and the unification of the colors so that the color continuity occurs. The basic determinants are: character's skin, sky blue, green grass and other objects or characters that should be uniform throughout the movie.

2.4 Main Drawbacks of the Linear Approach

Although the linear approach to making movies has been successfully used for decades, even in the top animated films of leading filmmakers such as Disney and Pixar, it has a number of disadvantages, through which the production of high quality films is very expensive and time-consuming. As a result, high-quality full-size animated films require very large resources, and budgets often exceed hundreds of millions of dollars. The most important factors that affect it are that each earlier phase must be completed before the next phase and the use of different tools in different phases. This means that mistakes made in earlier phases and detected in the next

Fig. 1 Refraction effects rendered with different ray bounce numbers [7]

phases are very expensive to repair. Creating animated films that are a kind of artistic work naturally require an iterative approach in which some imperfections or spots of necessary corrections are detected only when the final effect is seen. As discussed above, such modifications are extremely expensive to implement. An additional factor affecting the huge production costs in a linear approach is the use of tools that are very resource-consuming and require powerful computing power to achieve the required quality. This is related to very demanding algorithms in the field of lighting and rendering, such as ray-tracing, ambient occlusion where the quality and sense of realism increases with the number of iterations [6]. Depending on the number of iterations made, it may take several minutes, several hours or even several days to render a single frame of the movie. Since the target audience's expectations for top animated films are huge, manufacturers are doing everything possible to make the quality the best. An example of the influence of the number of samples for the refraction effect is shown in the Fig. 1, and a more detailed description of the algorithms for the example shown can be found in [7].

An example of the high-budget production of the animated film is "Coco" produced by Pixar Animation Studios and released by Walt Disney Pictures [8], in which the budget reached 175 million of dollars.

3 Non-linear Pipeline with Real-Time Game Engine

One of the solutions to the inconveniences of the traditional linear process is the use of a computer game engine. In this section, a general conception of the game engine is presented as well as the method of its use in the production of animation, as well as the advantages and disadvantages of using such a solution.

3.1 The Concept of the Game Engine

Originally, the game engines were designed to accelerate work on the production of 2D and 3D games, providing an integrated environment with re-usable components for graphics, animation, physics, music and special effects. Using the game engine, the developers did not have to implement everything from the scratch for each new game, and creating games became faster, cheaper and more profitable. In the first period, that is from the nineties to the first decade of this century, game engines were mainly used by programmers. Game assets such as character models, animations and special effects were created by artists, but rather with the use of external 3D tools, after which they were loaded from the source code of the game. Already at that time, graphic game editors were used, which were mainly used for simple tasks such as setting objects on the scene and possibly defining the places of the scene in which the player encountered certain events during the game (so-called Action Triggers). Everything, however, has changed dramatically in the last decade. Current game engines contain very rich game editors that are more artist-oriented. They also have integrated tools for modeling geometry, animation, special effects, physics or even procedural generation of the world [1, 2]. By the use of visual scripts, the game can be created entirely by designers and artists without writing a single line of code, as in the case of blueprints in Unreal Engine 4. The efficiency and flexibility of modern engines have made them used in areas other than computer games, everywhere where there is a need to use advanced 3D visualization with the potential need for interaction with the user. An example may be the production of animated films described here, as well as Augmented Reality, Television Broadcasting or various types of simulators.

3.2 Game Engines in Generating Animations

The main advantage of the game engine is the generation of a smooth sequence of high quality 3D frames in real time [9, 10]. In addition, it allows for easy change of the scene configuration (setting objects, textures, lights or cameras), where the view after changes is obtained almost immediately. These are very useful functions that eliminate many drawbacks of a linear production process.

Professional animation creators from leading companies differently identify the concept of real time in generating animation [3], namely as generating frames at a speed of at least 24 frames per second or as generating frames at a constant speed providing plausible realism. Another definition of real time allows lower frame-rate in generating frames (a few frames per second or even a few seconds per frame), but it allows for direct user interaction with playback or quick editing of the scene configuration. Regardless of the differences in real-time interpretation, it can be assumed that game engines are able to meet these assumptions.

The traditional linear pipeline of creating animated films contains many steps from the process of creating a concept to the final rendering. Optimizing this process with the use of a game engine does not eliminate these steps, but only changes the way they are implemented. Many tasks can be carried out simultaneously, and the improvement of those already completed can be performed without a large overhead on the whole production process. A an opposite to the traditional linear process, this approach will be called as non-linear one.

3.3 Non-linear Pipeline Benefits

The non-linear pipeline benefits are widely described in [3, 4, 11], and the most significant of them are listed below.

Interactive creative process with ease of editing and fast output Usually, all assets of the rendered scene have their representation directly in the engine, namely objects, figures, animations, lighting, special effects and elements of post-production. Artists can edit everything through a convenient graphical interface and get almost instant visualization of the result after the changes. This allows for quick iteration, and it is commonly accepted that more iterations usually lead to better results.

Earlier creative decision making In the general case, errors or corrections in the early phases (such as scenarios, storyboards or animatics) in a traditional linear pipeline are very expensive if they have to be handled in the later stages of production or post-production. In a non-linear pipeline, some of the work from pre-production and early production (e.g. animatics) can be implemented in the game engine, and the results are automatically applied in the later phases. In most cases, the pre-production and production phases can be carried out simultaneously. This causes that both artists and directors can see the almost final effect very quickly. This greatly facilitates key decisions and makes the entire production process more efficient.

Reduced production time Taking into account the aforementioned advantages of using engines, among others the possibility of parallel work of various departments over tasks from different stages of the process, the total production time is significantly shortened, which has a positive impact on the required budget of the project.

Reusable assets and brand consistency Very often, assets created in animated films are recreated for the needs of advertising materials, marketing, accompanying computer games, use on interactive websites, including digital television, Augmented Reality or Virtual Reality. Because game engines are currently used for most of these purposes, it is easy to create a consistent and integrated environment, using the same resources and ultimately to optimize the costs of the entire business around the movie.

3.4 Disadvantages of Using the Game Engine

The use of non-linear process and generation of frames in real time has many advantages over the traditional approach, but one of the most important drawbacks is still the quality. In the traditional approach with multi-million budgets, the most effective algorithms are used that give the highest possible quality without worrying about generating time. In real-time solutions, the most important factor is the speed of frame generation, which is why many simplifications are used in the algorithms used. This causes the need for a compromise between the quality of the result and the speed at which it was obtained. Thanks to the huge advances in the field of hardware and methods, the quality difference between approaches is constantly decreasing, but the fact is that it still exists.

4 Examples of Practical Applications

Although the concept of using game engines for serious animation production is relatively new, several films have been created, the quality of which can be compared with the best high-budget productions realized by companies such as Pixar and Disney. However, they have absorbed much less time and resources, and even relatively small production studios are able to implement them, provided they have the right skills.

4.1 GIFT

The "GIFT" project is one of the first attempts to use the game engine to produce high quality animated film [11]. For this purpose, the Unity engine [2] was used, which in this case was only a part of the process used mainly for the rendering of the configurable scene and the implementation of specialized shaders, leaving many elements of the traditional pipeline. However, it was enough to dramatically increase the efficiency of the process and the estimated reduction in production time by 20–30%. An exemplary film frame is shown in the Fig. 2 most on the left.

4.2 Fortnite

Epic company created a short animated film that was the teaser of one of their games, in which they used the Unreal Engine 4 engine. The whole movie consists of six sequences, each composed of 130 individual shots played at 24 frames per second. The production lasted 20 weeks and was carried out by a team of 5 to 20 people depending on the occupancy at the given stage. A detailed description of the process is presented in [3], and an exemplary film frame is shown in the Fig. 2 in the middle.

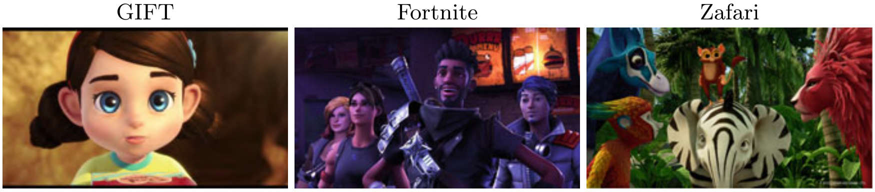


Fig. 2 Sample frames from movies rendered in real-time

4.3 Zafari

Zafari is the first animated television series created using the Unreal Engine 4 engine [12]. The aim was to create near-pixar look at television budgets and television speeds. The game engine was used here for lighting, effects and rendering, but the operation of the composition was not needed. The frames were generated at about 3 frames per second, and the process of generating a complete episode took from 3 to 4 h. The same result using a traditional farm render could take many days. An exemplary frame is shown in the Fig. 2 most on the right.

5 Discussion

Analyzing the current possibilities of film production transfer to real-time, it is worth discussing the most important aspects and challenges that have been identified by the representatives of the top companies that are involved in this revolution, including Pixar, Technicolor, Epic Games and Unity Technologies. This section describes the attempts of those experts to answer key questions raised during the recent SIGGRAPH 2018 conference in the field described [13].

What is the difference in making a game and making a movie? Generally, the process of "making pixels" is similar, but creating games is more of an engineering process and assumes a very large contribution of programmers, while the creation of films is mainly attended by artists. In addition, the process of creating a movie is more sequential, while in the process of creating games, the phases are largely carried out in parallel. The other difference is that after the production of the film, the product is actually finished, whereas the game requires a stabilization process (improvement of mistakes and satisfaction achieved by the players), which may be longer than the production process itself.

Why is real-time graphics important for making movies? The use of real-time in the production of films allows for more experiments and creative decisions in the development of the scenario, which in turn translates into a greater exploration of possibilities such as the use of physics and vehicles models used intensively in games. It also allows for closer contact of directors with experiments and results of intermediate phases, which ultimately leads to faster achievement of the expected results.

Does real-time playback make sense in the movie world? The film is mainly a storytelling which can take different forms. For future generations, the interaction may be an indispensable element in storytelling, which will make playback in real-time necessary. In addition, the filmmakers themselves see a lot of interesting impressions for the viewer in interacting with the movie (e.g. watching a movie on the iPad and changing the camera view). On the other hand, the director decides how to show the story to the viewer. When the viewer has the opportunity to change the way the director has planned, there is a risk that the storyteller cannot tell the whole story.

Another aspect is the use of stories in alternative media such as broadcasted TV shows, live action events and other ways for entertaining people in which real-time playback is essential.

How far the game technology is helpful for film making? Currently, games and film productions are developing independently, and attempts to converge them should theoretically lead to a certain breakthrough in which they can “explode” and expand their capabilities. It is difficult to imagine now all the possibilities that will appear, because at the moment, apart from a few attempts, these two threads are still in parallel.

Is real-time adaptation a software, hardware or process problem? Experts agree that the adaptation of film production to real time applies to all: software, hardware and process. The current quality of graphics in real time, similar to the quality obtained in a traditional pipeline, would not be possible without the evolution of NVIDIA cards, such as used in the DGX station [14]. The availability of software, which in many cases is free and is available to everyone, means there is no limit to do “cool stuff”. On the other hand, you can not be under the delusion that the real-time process will solve all problems and will not cause new ones, so the process should be changed slowly and carefully.

Does the real-time really saves the budget If someone can quickly see something similar to the final result, as is the case with real-time, it certainly affects the money savings. It should be noted, however, that artists tend to do things as well as possible. So with real-time technology, they usually use it to do more iterations, more iterations usually provide better quality. In this case, the game engine as optimization improves the quality of the result, not necessarily reducing the time of work and reducing costs.

How real time helps small movie makers? Easy access to real-time technology makes it easier to create high-quality movies by smaller creators. However, there are still classic barriers: everyone can make a simple movie, but not everyone is able to make a product at the level of top directors in the world.

What are the biggest challenges of using real-time in making movies? One of the most difficult tasks is to achieve convergence between the results of the traditional pipeline and the one based on real time. For example, how to obtain the quality of ray-casting from a traditional pipeline in real-time, where some other methods are used. Of course, the results may be slightly different, they may be noisy but they should be very similar. As there are more and more hardware and software solutions, another important challenge is also the selection of the right and the most optimal direction to achieve maximum accuracy and realism in real time. Another challenge is visual parity in assets: how to get similar quality of the same asset in both traditional and real time. For example, for assets in the traditional pipeline, there is no penalty for too big textures, while in games the speed important, so the Level of Details (LOD) is used. The important aspect is the ease of adaptation of game technology tools to various film production pipelines, which may differ significantly among creators. Regarding compositing, current engines do not have sufficiently ordered methods, because they are usually not used in compositing games. It is a big field for research and improvement.

Will be two versions of the same asset needed? One can imagine that there will be a need to use different versions of the asset (one for the offline rendering and the other for realtime), differing in the amount of details. However, considering the increasingly efficient game engines, the difference between the details handled in the offline rendering and real-time rendering is getting smaller. For this reason, it can be anticipated that the same asset with a large amount of detail will be supported in both cases.

What are the top points in which the game engine can help? As a summary, the most important factors are following: quick decision making, a quick look at large environments for decision makers, easy experiments for all project members as a giant window for creativity.

Is the future of film making in real time? In general, the basic purpose of the film is story telling. Of course, the artists choose their own preferred techniques of story telling and the current generation of creators strongly rely on proven traditional methods. But at the same time a new generation is coming, who will want to do it in a different way. Current young people, for the moment mostly the recipient of content, expect greater customization and some sort of gamification in the reception of both films and television on various devices (e.g. on mobile phones). Here, in turn, real-time game engines seem to be the optimal solution.

6 Conclusion

We have presented the traditional approach to the production of animated films used by top companies like Pixar and Disney, as well as pointed out the inconvenience and factors that make creating high-quality animations require large budgets. We have also described the new approach of this process by using game engines in which real-time rendering is to optimize the traditional pipeline, we pointed out the advantages and disadvantages of such an approach, we showed recent attempts to use it in practice, and we described its most important aspects and challenges for the future.

It turns out that the development of hardware technology and software tools, mainly in the form of modern game engines oriented also to artists, has opened the possibility of creating high quality animated films with much smaller budgets than in traditional approaches. Although there are still slight differences in quality, existing tests confirm that the transformation of animation pipelines towards real-time can introduce savings in both time and money.

Considering also the current tendency for convergence of various branches of entertainment, such as the use of the same assets in films, TV broadcast and games, the nearest trend in film making is to be real-time, interactive and non-linear.

Acknowledgements This work has been supported by the National Centre for Research and Development, Poland, in the frame of Operational Programme Smart Growth 2014–2020 within Sectoral Programme GAMEINN, POIR.01.02.00-00-0115/16 “Innovative use of computer game engine in order to reduce costs and time of production of animated film”.

References

1. Epic Games: Unreal Engine 4 Documentation (2018). https://docs.unrealengine.com/en-us/
2. Unity Technologies: Unity Manual (2018). https://docs.unity3d.com
3. Pohl, B., Brakensiek, T., Lombardo, S.: Fortnite Trailer: Developing a Next Gen pipeline for a faster workflow. In: Epic Games white Papers (2018)
4. Epic games: why real-time technology is the future of film and television production. In: Epic Games white Papers (2017)
5. Dunlop, R.: Production Pipeline Fundamentals for Film and Games. Focal Press (2014)
6. Ritschel, T., Dachsbacher, C., Grosch, T., Kautz. J.: The state of the art in interactive global illumination. Comput. Graph. Forum **31**(1), 160–188 (2012)
7. Keyshot 3 User's Manual (2012). https://www.keyshot.com/keyshot3/manual
8. Pixar Feature Films: Coco (2017). https://www.pixar.com/feature-films/coco
9. Johnson, P., Pettit, D.: Machinima: The Art and Practice of Virtual Filmmaking, McFarland (2012)
10. Marino, P.:, 3D Game-Based Filmmaking: The Art of Machinima, Creating Animated Films with 3D Game Technology, Paraglyph (2009)
11. Sanchez, M.: Progressing Beyond Pre-Render: The MARZA Movie Pipeline for Unity (2016). https://blogs.unity3d.com/2016/09/06/progressing-beyond-pre-render-the-marza-movie-pipeline-for-unity
12. Failes, I.: Upcoming Animated Series "Zafari" Is Being Rendered Completely With The Unreal Game Engine (2017). https://www.cartoonbrew.com/tools/upcoming-animated-series-zafari-rendered-completely-unreal-game-engine-153123.html
13. Jeremias-Vila, P., Libreri, K., Quaroni, G., Tatarchuk, N., Fagnou, D.: The present and future of real-time graphics in film. In: ACM SIGGRAPH 2018 Panels, articleno. 5, pp. 5:1–5:2 (2018)
14. NVIDIA DGX Station Personal AI Supercomputer (2018). https://www.nvidia.com/dgx-station

Visual Keyboards or QR Codes in an Inclusive School Environment

Vojtěch Gybas, Libor Klubal and Kateřina Kostolányová

Abstract This post is based on previous world studies on the use of iPads in teaching pupils with special educational needs. Previous studies have concluded that the iPad's visual keyboard reduces the pupil's graphical load (compared to the traditional keyboard of the computer) and positively compensates for partial deficits in areas of fine motoring. In the form of a pedagogical experiment, we tried to show the impact of the visual keyboard and QR codes on the work of pupils with an Internet browser in the inclusive school environment. What is more convenient for a pupil—working with a visual keyboard or working with QR codes and scanning the destination Internet address?

Keywords iPad · Safari browser · QR codes · Pupil with special educational needs

1 Introduction

This paper seeks to respond to the current trends in education in the Czech Republic for pupils in the inclusive environment of ordinary elementary schools, respecting modern ICT. Using iPads supports the development of literacy skills, provides easy access to information and is much easier for students with disabilities [1]; using iPads has helped reduce behavioral problems [2]; while iPads are considered socially acceptable, which is less stigmatizing for students with special needs [3]. Using iPads enables teachers to reinforce classroom learning and engage reluctant pupils in learning processes [4] while developing both weaker and stronger pupils together [5]. Teachers and pupils have quickly found that using iPad is a motivating and effective learning instruction tool. There are also innovative iPad apps that facilitate learning and learning for mentally handicapped pupils and demonstrate a greater degree of independence and positive behavior in the pupils' educational curriculum [6, 7].

V. Gybas (✉) · L. Klubal · K. Kostolányová
University of Ostrava, Fráni Šrámka 3, Mariásnké Hory, Ostrava, Czech Republic
e-mail: vojtech.gybas@osu.cz

M. Huk et al. (eds.), *Intelligent Information and Database Systems: Recent Developments*, Studies in Computational Intelligence 830,
https://doi.org/10.1007/978-3-030-14132-5_17

Waddington et al. [8] reported in his study focusing on using iPads with three pupils with Autistic Spectrum Disorder that the iPad can be used as a form of communication tool for pupils with limited or undeveloped speech.

Writing on the iPad requires a lesser load on grammoto and facilitates visual and sensory learning [9]. The vogue concept of writing in the air [10] is relevant here because we consider the iPad to be a new cultural tool that offers a different kind of work in space. With iPads, on-screen typing can be converted from a digitally created artifact to a physical artifact screen on paper, meaning writing is instantly visible and can also be printed in another tangible form in a few moments. We have confirmed the benefits of iPads even in of our partial research, when we reached concluding that the use of the iPad is more appropriate in view of its visual keyboard comparison with ordinary PCs both for pupils with special educational needs and for seniors. It is obvious that the iPad and the visual keyboard affect both the whole process of learning and the reduction in pupil's grammoto. As is the case when the teacher wants to make work easier for pupils by creating QR codes. Does that make any sense?

2 Research Design

Pupils with special educational needs in the inclusive environment of the elementary school participated in a pedagogical experiment that tried to prove the impact of the visual keyboard relative to the pupil's graphomotor and at the same time to prove what is better for the teacher from his own view. Specifically, whether working with a student's visual keyboard (input of web address text) is more convenient for pupils and puts less on their grammatical skills, or it is easier for pupils to work with read the QR code reader and scan your destination URL. What makes pupils more comfortable? And what better for a teacher? Create QR codes or let the pupils work freely?

2.1 Characteristics of the Research—A Pupil at a Special Elementary School

29 pupils participated in the pedagogical experiment, 7th year (age 13), of which 6 pupils with special educational needs. All 6 pupils have IQ ranging from 50 to 69 points. These pupils have a slight mental retardation (LMR), the remaining pupils are in the normal IQ band. Within the education system, these 6 pupils are included in the so-called 3rd level of support measure. The experiment was on the first and last day of the week, Monday and Friday. Every time in Informatics. Experiment URLs were:

- https://www.seznam.cz/
- https://cs.wikipedia.org/wiki/IPad
- http://androidmarket.cz/ruzne/historicky-vyvoj-tabletu-infografika/.

All pupils in the class work according to their individual skills with iPads and with educational applications. During the classroom, difficulties often arise when the teacher asks for a class to get a class in The iPad has opened a web page. In there is a problem for a while. Pupils ask for address, cancel, make mistakes from inattention. Complete assignment one instruction to opening a website takes too much time. Pupils were continuously acquainted with Safari apps and way of entering a web address, were also introduced the option to load a QR code through the Camera application, when you then need to tap your finger on the dialog box, then a web page appears.

3 Working with Visual Keyboard and Safari Browser

WWW address 1: www.seznam.cz—first pupils were invited to input of the first web address. The assignment from the teacher was to open the "list.cz" page via the Safari web browser, and the lecturer wrote it on the blackboard. Pupils must unlock iPad, launch Safari, and open a web page. Teacher with the assistant subsequently measured the times of individual pupils. Measurements were run on the iPad, at application Clock, form: split times. The resulting time for individual pupils was then counted.

IN first group (pupils without LMR)—4 pupils wrote "list", then entered "open" and then selected from "seznam.cz" offers. 19 pupils wrote straight "seznam.cz" or "www.seznam.cz". IN of the second group (pupils with LMR)—4 entered "list", then "open". Then they chose from "seznam.cz" offers. 2 pupils had problems, they entered the wrong word. Instead of the "list" they entered the "list", respectively. "Zesnam". Therefore, in could not find the address. The teacher had to help these two students. The opening of this website had only one goal. Verify pupils' readiness for the job, as this was only the home page of the Czech web portal (Fig. 1).

WWW address 2: https://cs.wikipedia.org/wiki/IPad—When you enter this URL, pupils are advised to carefully copy the text. Only then will the web page appear properly. The teacher wrote the web page again on the blackboard. Even though the URL was written on the whiteboard, the pupils had questions: what was the address? How did you say it?, etc. It was obvious that the address was already more complex.

The first group (pupils without LMR)—5 pupils knew the wikipedia portal. That's why they wrote wikipedia.cz, then searched for the word iPad. 18 students scanned the whole address. They were slower than a group of 5 pupils.

The second group (pupils with LMR)—there was already a greater time load, especially with by typing ":", "/". Manipulating with individual characters make the pupils the biggest problem. The address eventually managed everyone, but the first group had to wait for them (Fig. 2).

	21-25s	26-30s	31-35s	36-40s	41-45s	46s and more
The first group	1	7	7	3	5	0
The second group	0	2	2	0	0	2

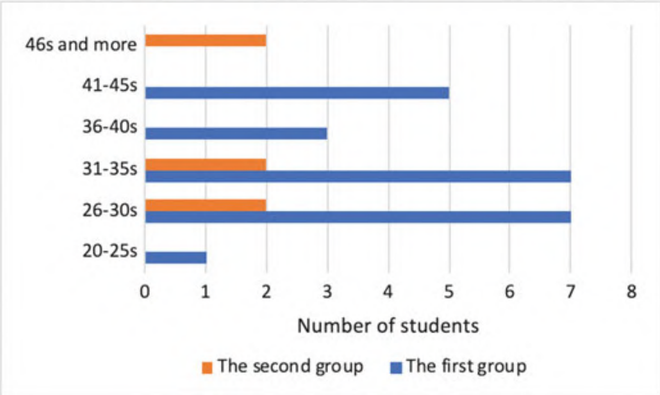


Fig. 1 The speed of opening 1. Web addresses

	1-5s	6-10s	11-15s	16-20s	21s and more
The first group	17	2	2	0	0
The second group	0	0	4	2	0

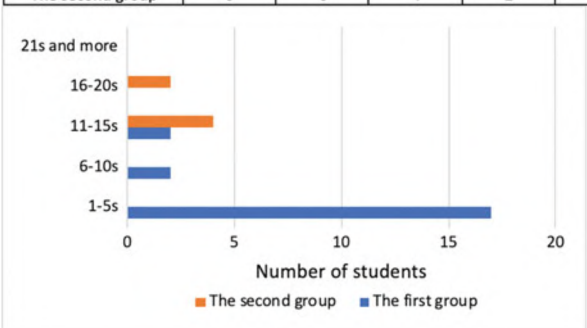


Fig. 4 The speed of opening 1. URLs using QR code

WWW address 3: http://androidmarket.cz/ruzne/historicky-vyvoj-tabletu-infografika/—again the same procedure. The teacher said the address they were looking for and wrote it on the blackboard.

The first group did not know the address. They could not guess what they could write and what to look for, so it was obvious that all 23 students fairly described the entire address. As a result, the resulting time was based only on how quickly the pupils managed to write the address.

The second group—there was a longtime lag. Switching the keyboard to characters and back to letters has been a problem for 4 pupils who eventually managed to address

	21-25s	26-30s	31-35s	36-40s	41-45s	46s and more
The first group	0	5	11	7	0	0
The second group	0	0	0	0	1	5

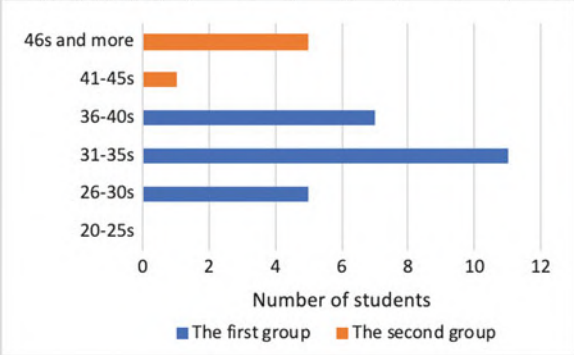


Fig. 2 Speed of opening 2. Web addresses

	21-25s	26-30s	31-35s	36-40s	41-45s	46s and more
The first group	0	1	2	1	3	16
The second group	0	0	0	0	0	6

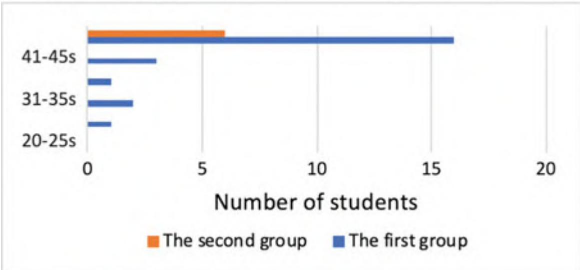


Fig. 3 Speed of opening 3. Web addresses

the address with using a teacher assistant. 2 pupils managed to copy the address themselves (Fig. 3).

4 Work with QR Codes and Camera App

The teacher will display the first QR code—www.seznam.cz and ask the pupils to scan it for the iPad, and the pupils have to display the webpage that the code will offer.

The first group—21 pupils manages independently, without any problems and very quickly. 2 pupils instead of scanning the code took a photo, so they pressed the

	1-5s	6-10s	11-15s	16-20s	21s and more
The first group	23	0	0	0	0
The second group	4	0	2	0	0

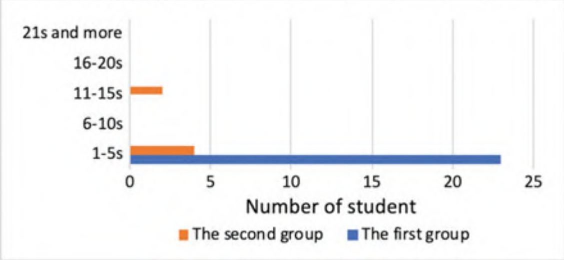


Fig. 5 Speed of opening 2. URLs using QR code

	1-5s	6-10s	11-15s	16-20s	21s and more
The first group	23	0	0	0	0
The second group	6	0	0	0	0

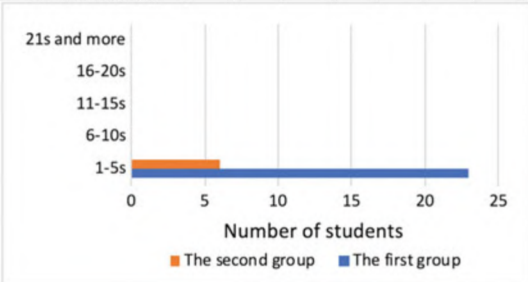


Fig. 6 Opening speed 3. URLs using QR code

camera's shutter button. They then ask the teachers that nothing is displayed. Teacher helps, restarts the Camera application again. Pupils now do not miss it.

The second group—all 6 pupils took a photo. So, no webpage appeared. Teacher and teacher assistant respond very promptly, launching Camera again. This is all without a problem and very quickly this time. The resulting time is higher (Fig. 4).

Teacher now displays the second QR code for the https://cs.wikipedia.org/wiki/IPad webpage and asks the students for the same task.

The first group—again scans, all pupils without any problem and very quickly.

The second group—4 pupils scan. Seamlessly. 2 pupils are asked whether or not again in the same way as in previous case. The teacher responds and answers. Pupils scan. Seamlessly. But due to the query with higher time lag (Fig. 5).

Teacher Shows QR Code with web site http://androidmarket.com/ruzne/historicky-vyvoj-tabletuinfografika/ and asks students to scan.

The first group—scans, without any problems, very quickly.

The second group scans, without any problem or query, very quickly (Fig. 6).

5 Conclusion

From experimental investigation, we can assume that working with QR codes reduces the load on the pupil's graph. Compared to working on the visual keyboard, it's clear that QR code scans are faster than writing the URL itself. There are no unnecessary organizational problems when the pupil does not understand the task. QR codes simply scan and continue to work. Pupils fulfill the assignment and the teacher does not have to be burdened with unnecessary things. Scanning can certainly be of great benefit to enrolled pupils who may have reduced intellectual efficiency. These pupils in Inclusive environments usually have the opportunity to use the assistant teacher's assistance in teaching. In such a structured teaching, pupils become more self-sufficient, less mistaken in their tasks, naturally they are self-empowered and experience a higher sense of success.

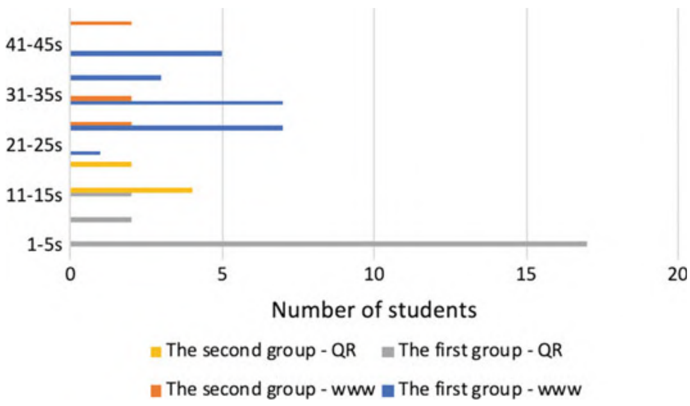


Fig. 7 Comparison of job speed 1: writing text versus QR code

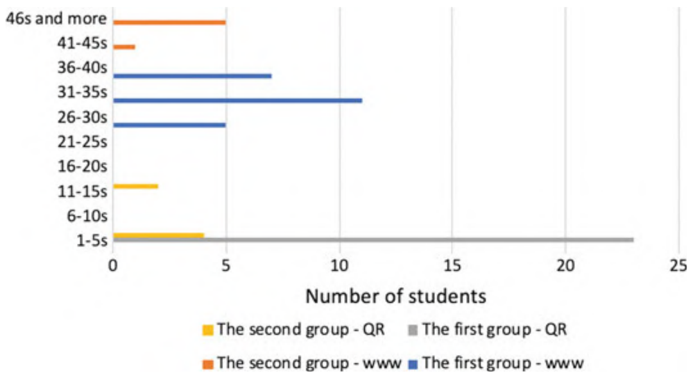


Fig. 8 Comparison of job speed 2: writing text versus QR code

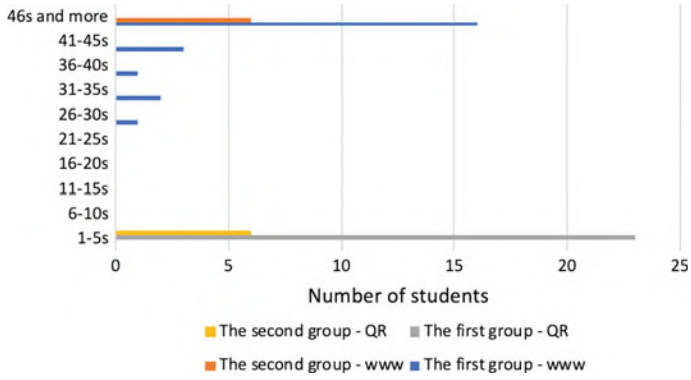


Fig. 9 Comparison of job speed 3: writing text versus QR code

For pedagogues, QR codes should be created by motivation. One such code that contains a web page will create in 2–3 min. This time he will return to him the form of teaching and the almost uniform working pace of all pupils. Already without distinction of pupils' individual peculiarities (Figs. 7, 8 and 9).

References

1. Spooner, F., Ahlgrim-Delzell, L., Kemp-Inman, A., Wood, L.A.: Using an iPad2® with systematic instruction to teach shared stories for elementary-aged students with autism. Res. Pract. Persons Severe Disabil. **39**(1), 30–46 (2014)
2. Neely, L., Rispoli, M., Camargo, S., Davis, H., Boles, M.: The effect of instructional use of an iPad[R] on challenging behavior and academic engagement for two students with autism. Res. Autism Spectr. Disord. **7**(4), 509–516 (2013)
3. Kagohara, D.M., van der Meer, L., Ramdoss, S., O'Reilly, M.F., Lancioni, G.E., Davis, T.N., et al.: Using iPods(r) and iPads(r) in teaching programs for individuals with developmental disabilities: a systematic review. Res. Dev. Disabil. **34**, 147–156 (2013). https://doi.org/10.1016/j.ridd.2012.07.027
4. Gentile, M.: The importance of managing iPads in the classroom. Educ. Digest **78**(3), 11–13 (2012). Retrieved from www.ed.com
5. Foulger, T.S., Walker, M.L., Burke, D., Hansen, R., Willimas, M.K., Slykhuis, D.A.: Innovators in teacher education: diffusing mobile technologies in teacher preparation curriculum. J. Digital Learn. Teacher Educ. **30**(1), 21–29 (2013)
6. Cumming, T., Rodríguez, C.: Integrating the iPad into language arts instruction for students with disabilities: engagement and perspectives. J. Spec. Educ. Technol. **28**(4), 244–250 (2013)
7. Rodríguez, C.D., Strnadová, I., Cumming, T.: Using iPads with students with disabilities. Intervention Sch. Clin. **49**(4), 244–250 (2014)
8. Waddington, H., Sigafoos, J., Lancioni, G.E., O'reily, M.F., van der Meer, L., Carnett, A., Stevens, M., Roche, L., Hodis, F., Green, V.A., Sutherland, D., Lang, R., Marschik, P.B.: Three children with autism spectrum disorder learn to perform a three-step communication sequence using an iPad®-based speech-generating device. Int. J. Dev. Neurosci. **39**, 59–67 (2014). ISSN 0736-5748 [on-line]. Cit. Dne 30. 9. 2018. Dostupné na https://doi.org/10.1016/j.ijdevneu.2014.05.001

9. Flewitt, R., Kucirkova, N., Messer, D.: New directions for early literacy in a digital age: the iPad. J. Early Child. Literacy. Published online 20 May 2014. https://doi.org/10.1177/1468798414533560
10. Vygotsky, L.S.: Mind in Society: The Development of Higher Psychological Processes. Harvard University Presss, Cambridge, MA (1978)

Human Identification Based on Shallow Learning Using Facial Features

Van-Dung Hoang, Cong-Hieu Le and The-Anh Pham

Abstract Today, identity recognition systems have achieved high accuracy and widely used in specific application areas such as recognition system based on retina imaging in immigration inspection, civil security and citizen management. Identity recognition is a very important task in intelligent surveillance systems. In these systems, human is required to be submissive for data acquisition to identify themselves. However, the automated monitoring systems are required to be active for information retrieval and human is passively monitored in this situation. In this kind of approach, human recognition is still a challenging task for the overall system performance. This study proposes a solution for human identification based on the human face recognition in images extracted from conventional cameras at a low resolution and quality. Our proposed approach for human identification is based on histogram of oriented gradients (HOG) feature descriptor and Support vector machine (SVM) classifier using a similarity matric estimation. The proposed method was evaluated on some standard databases which are available online and on our own collected dataset.

Keywords Personal identification · Face features · Intelligent monitoring systems

1 Introduction

In recent years, with the rapid development of science and technology, monitoring systems based on artificial intelligence techniques have been studied and applied

V.-D. Hoang (✉)
Quang Binh University, Quang Binh, Vietnam
e-mail: dunghvqb@gmail.com

C.-H. Le
Electric Power Department, Quang Tri, Vietnam
e-mail: hieu.lecong@gmail.com

T.-A. Pham
Hong Duc University, Thanh Hoa, Vietnam
e-mail: phamtheanh@hdu.edu.vn

M. Huk et al. (eds.), *Intelligent Information and Database Systems: Recent Developments*, Studies in Computational Intelligence 830,
https://doi.org/10.1007/978-3-030-14132-5_18

in many fields of security monitoring, medicine application, education, intelligence transportation systems, etc. [1–5]. The data should be received within a passive manner in applications such as security monitoring systems, detecting crime in the crowds and so on. In the first kind of the applications, data of human retina, fingerprint can only be received when human is required to provide information for the systems. However, there are some surveillance systems which are required to be actively and fully automated in the task of data acquisition, meanwhile recognition systems based on retinal image, fingerprints could not be done. In this kind of applications, a human recognition system based on face, or human walking style information is applicable. The method is a cascaded multimodal biometric system with combination of fingerprint and iris recognition. The key ideas for fingerprint iris recognition are minutiae extraction of fingerprint information and encoding the Log Gabor filtering. The experiments on some dataset demonstrated the superior performance of that approach compared to unimodal systems. Another method also used iris for biometric identification presented in a pattern [6]. Meanwhile authors in [7] presented a technique for online iris image compression and personal identification, which called iris image compressor and identifier. That method converts the iris image of eye in the form of Laplace–Beltrami Spectra, then Laplace–Beltrami spectra is converted into the form of Strakos matrix. The matrices of Eigen values are calculated for identifying a person. The proposed method was experimented on a hundred iris images of CASIA database. The experimental results illustrated that the method is robustness, efficient and economically feasible. In another paper, Patwari et al. [8] presented an algorithm using retinal blood vessels bifurcation for personal identification. Their method focuses on detection of retina blood vessels and measurement bifurcation points of blood vessels. The proposed method was performed on real data of 300 images which collected and analyzed by doctors. That algorithm achieved high accuracy score, which is appropriate for real application.

The standard approach investigated HOGs for human feature description in [9, 10]. To deal with partial occlusion, [11, 12] combined HOG and Local Binary Pattern (LBP) for feature description. In that system, the authors accumulated both HOG and LBP to construct feature vectors, which are fed to the SVM algorithm in both the training and detection stages. Experiments indicated that the system was capable of handling partial occlusion. With some modifications of the feature descriptors, the human detection system in [13] was built with an extended histogram of gradients as a feature; it uses an unsigned histogram of gradients and a histogram of absolute values of opposite directional in a cell. Experiment results showed that the system based on an extended histogram of gradient outperformed the standard HOG. Systems based on HOG features descriptors and SVM learning achieve significant accuracy in human detection. However, they are usually high in computational cost [14].

The research provides a uniqueness assessment of palatal rugae based on their anatomical 3D conformations. Experimental results showed that the method achieved high precision and applicable. A personal recognition method based on hand back skin texture is proposed in [15]. In that paper, authors investigated problem of the hand back skin texture pattern for personal identification and gender classification. The model was trained and classified the hand back skin texture patterns in good result scores. Evaluation results demonstrated an efficient for assist human identification and gender classification. Meanwhile, Yang et al. [16] proposed an approach based on fusion at the feature level of finger vein and finger dorsal texture. The method was compared the state-of-the-art unimodal biometrics on the performance. Compared results on established dataset illustrated that their approach achieved higher identification accuracy and lower equal error rates to other methods. In different approach, our contribution presents an identity recognition approach based on deep learning techniques using facial image data, which is collected from surveillance camera with normally quality of image.

2 Proposed Method

Human identification plays an important role in many applications such as monitoring and human management systems. This research focuses on identity recognition based on face information which is extracted from surveillance cameras. According

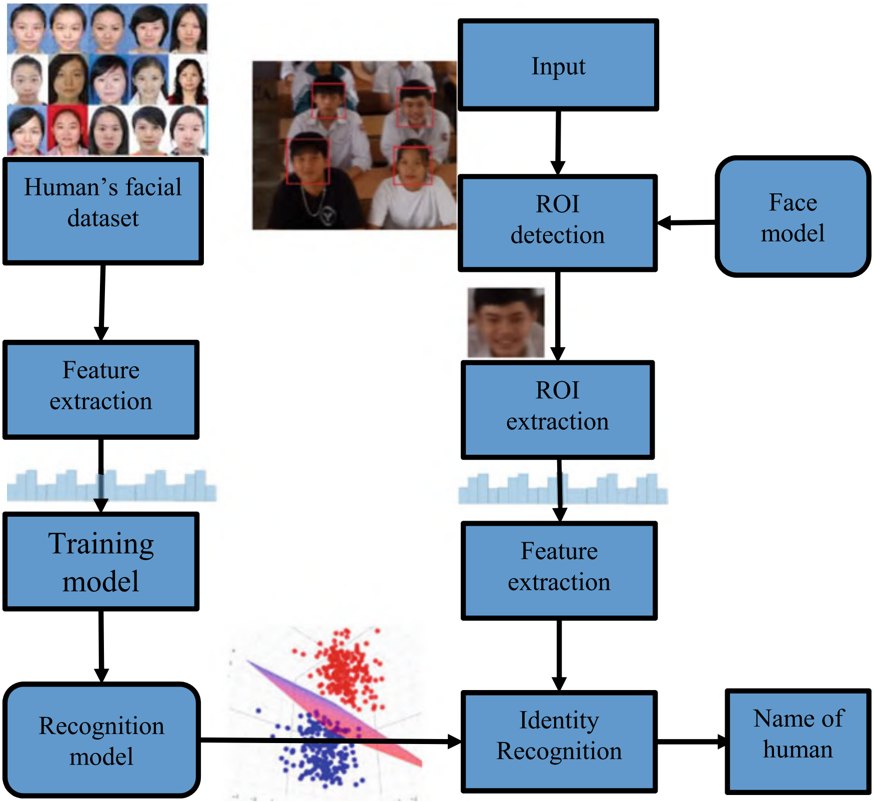


Fig. 1 Flowchart of the identity recognition system

to previous discussion, identity recognition uses special information of a person such as retinal, irises, fingerprints, and other types of biometrics often provide high accuracy and have been applied in practical systems. However, there are some types of applications that cannot coerce human to provide their information. In this situation, the surveillance systems have to be active for collecting the data and human is to be passive in the process of systems. This paper presents some results on identity recognition task studied by the use of artificial intelligence and image processing techniques for face detection and identification. The general scheme of the identity recognition system is shown in the diagram in Fig. 1.

In the detection task, a facial model was trained on a face dataset using the Haar-like feature and AdaBoost classifier, resulting in a detection model with detection rate of 98.5%. We provide a recognition solution based on the tasks of feature extraction and identity recognition using SVM approach instead of deep learning. There is a number of approaches for this problem with the use of shallow learning techniques based on traditional a feature extraction method. In traditional approaches, there are many methods for feature extraction such as HOG, SHIFT, LBP, Haar-wave…. The feature extraction results are then fed as inputs for classifiers trained by SVM, boosting, decision tree, GA, etc. In our approach, histogram of oriented gradients feature descriptor and Support vector machine is used for both the objectives of feature extraction and classification task.

3 Experimental Results

We evaluated the proposed method on some datasets as a basis for the evaluation. The proposed method is carried out by the traditional method based on the use of HOG feature descriptor and SVM identification. Moreover, in some situations, shallow machine-based human identification is more suitable for applications than deep learning [5]. In the following subsections, details of the results evaluated on several different datasets will be presented.

3.1 Staffhome Dataset

The Staffhome data set was created by the research team [17]. Face images are taken by looking directly at different lighting conditions. The images are cut and standardized so that the image only contains enough faces. Images was created for the purpose of facial recognition and reconstruction. The image data set is available on their website. The Staffhome dataset includes 3494 faces, taken from 84 people. Each person takes a sample of about 24–48 face images. Some face patterns are shown in Fig. 2. Test results are shown in Table 1 and Fig. 3.

Fig. 2 Some facial examples from Staffhome dataset

Table 1 Experimental results on Staffhome dataset

	HOG + SVM	Deep CNN
Number samples for training	2101	2101
Number samples for testing	1393	1393
Computational time for training feature extraction (s)	73.19	8654
Computational time for train (s)	926.22	
Computational time for test feature extraction (s)	43.02	730
Computational time for test (s)	450.30	
Accuracy (%)	98.84	98.99

3.2 Our Dataset

The experimental dataset was collected by our team, which is called Student-1 dataset. The data samples were taken by a low-resolution camera which mounted in a class of a high school. Facial images were captured under different lighting conditions. Then sample images were cropped and normalized so that each image contains the entire head with a full face. The Student-1 dataset contains 5083 faces of 30 different students. Each student was sampled from 100 to 200 facial samples. Some face patterns are shown in Fig. 4. In our experiment, we used about 60% samples of each student for training and about 40% sample of each student for evaluation (Fig. 5; Tables 2 and 3).

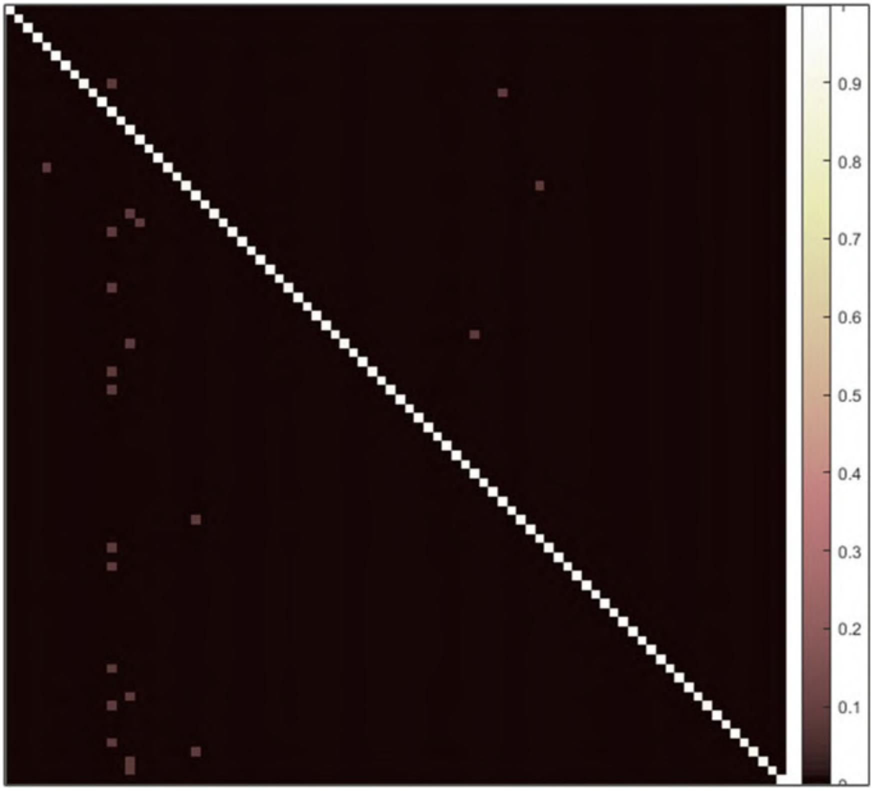


Fig. 3 Confusion matrix of recognition rate on Staffhome dataset

Table 2 Details of facial dataset used for training and evaluation

	#1	**#2**	**#3**	**#4**	**#5**	**#6**	**#7**	**#8**	**#9**	**#10**	
# Training samples	98	107	105	90	106	106	103	104	106	94	
# Test samples	66	71	70	60	71	70	68	70	71	62	
# All samples	164	178	175	150	177	176	171	174	177	156	
	#11	**#12**	**#13**	**#14**	**#15**	**#16**	**#17**	**#18**	**#19**	**#20**	
# Training samples	104	113	82	102	93	91	55	79	106	109	
# Test samples	69	76	55	68	62	61	36	53	71	72	
# All samples	173	189	137	170	155	152	91	132	177	181	
	#21	**#22**	**#23**	**#24**	**#25**	**#26**	**#27**	**#28**	**#29**	**#30**	**#31**
# Training samples	107	102	103	109	85	90	85	107	94	65	151
# Test samples	71	68	69	73	56	60	56	71	63	43	100
# All samples	178	170	172	182	141	150	141	178	157	108	251

Fig. 4 Some facial examples of our collected dataset

Table 3 Experimental results on our collected dataset

	HOG + SVM	Deep CNN
Number samples for training	3051	3051
Number samples for testing	2032	2032
Computational time for training feature extraction (s)	51.55	21,351
Computational time for train (s)	121.11	
Computational time for test feature extraction (s)	29.30	156
Computational time for test (s)	32.04	
Accuracy (%)	93.18	94.39

4 Conclusions

In this study, we proposed an approach based on traditional methods of HOG and SVM identity human recognition. Experimental results showed that the recognition system based on our proposed method for human identification using facial images reached acceptable accuracy rate. Our proposed method has been studied on the basis of empirical data, comparing of several machine learning methods to provide an effective solution for people identification problem. Our contribution consists of conducting evaluations to help the selection of an effective method for

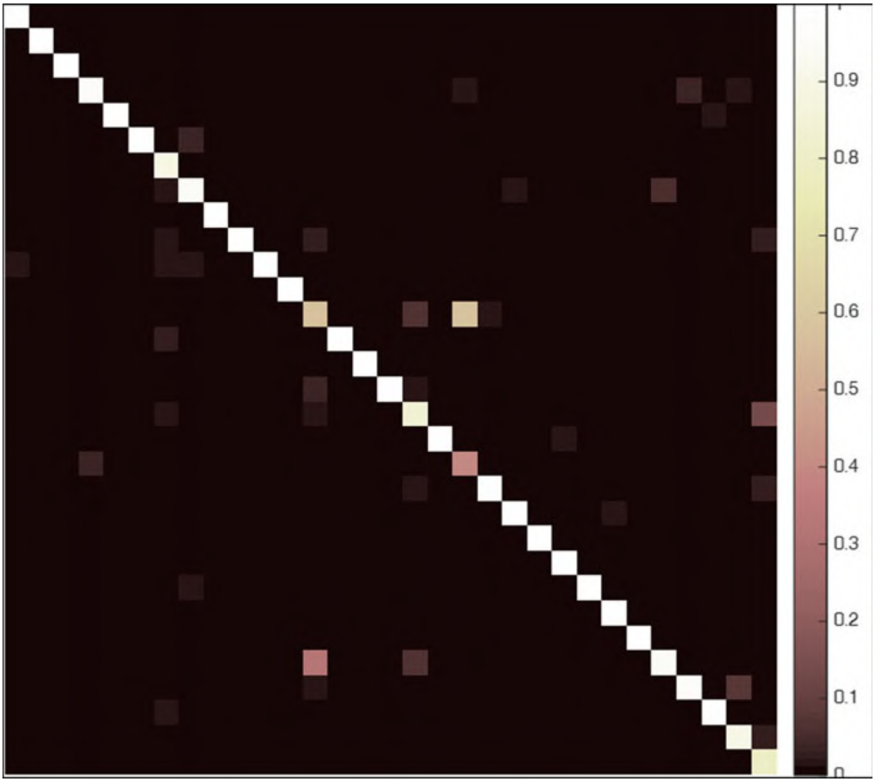


Fig. 5 Confusion matrix of recognition rate on our collected dataset

human identification problem and providing resulted details of each person identity with high accuracy to compared to deep learning based on convolutional neural network. Recently, we are improving the recognition model by experimenting on bigger dataset, so that the system can be applied to real application in intelligent monitoring systems.

References

1. Hattori, H., Lee, N., Boddeti, V.N., Beainy, F., Kitani, K.M., Kanade, T.: Synthesizing a scene-specific pedestrian detector and pose estimator for static video surveillance. Int. J. Comput. Vis. 1–18 (2018)
2. Sindagi, V.A., Patel, V.M.: A survey of recent advances in cnn-based single image crowd counting and density estimation. Pattern Recogn. Lett. **107**, 3–16 (2018)
3. Hoang, V.-D., Jo, K.-H.: Path planning for autonomous vehicle based on heuristic searching using online images. Vietnam J. Comput. Sci. 1–12 (2014)

4. Pham, T.-C., Luong, C.-M., Visani, M., Hoang, V.-D.: Deep CNN and data augmentation for skin lesion classification. In: Asian Conference on Intelligent Information and Database Systems, pp. 573–582. Springer, Berlin (2018)
5. Tran, D.-P., Hoang, V.-D., Pham, T.-C., Luong, C.-M.: Pedestrian activity prediction based on semantic segmentation and hybrid of machines. J. Comput. Sci. Cybern. **34**, 113–125 (2018)
6. Hanna, K.J.: System and method for iris data acquisition for biometric identification. Google Patents (2018)
7. Mishra, K.N.: An efficient technique for online iris image compression and personal identification. In: Proceedings of International Conference on Recent Advancement on Computer and Communication, pp. 335–343. Springer, Berlin (2018)
8. Patwari, M.B., Manza, R.R., Rajput, Y.M., Saswade, M., Deshpande, N.: Personal identification algorithm based on retinal blood vessels bifurcation. In: 2014 International Conference on Intelligent Computing Applications (ICICA), pp. 203–207. IEEE (2014)
9. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: Conference on Computer Vision and Pattern Recognition, pp. 886–893 (2005)
10. Hoang, V.-D., Jo, K.-H.: Joint components based pedestrian detection in crowded scenes using extended feature descriptors. Neurocomputing **188**, 139–150 (2016)
11. Wang, X., Han, T.X., Yan, S.: An HOG-LBP human detector with partial occlusion handling. In: International Conference on Computer Vision, pp. 32–39 (2009)
12. Won-Jae, P., Dae-Hwan, K., Suryanto, S., Chun-Gi, L., Tae Moon, R., Sung-Jea, K.: Fast human detection using selective block-based HOG-LBP. In: 2012 19th IEEE International Conference on Image Processing (ICIP), pp. 601–604 (2012)
13. Satpathy, A., Xudong, J., How-Lung, E.: Extended histogram of gradients feature for human detection. In: IEEE International Conference on Image Processing, pp. 3473–3476 (2010)
14. Enzweiler, M., Gavrila, D.M.: Monocular pedestrian detection: survey and experiments. IEEE Trans. Pattern Anal. Mach. Intell. **31**, 2179–2195 (2009)
15. Zhang, D., Lu, G., Zhang, L.: Hand back skin texture for personal identification. In: Advanced Biometrics, pp. 213–233. Springer, Berlin (2018)
16. Yang, W., Huang, X., Zhou, F., Liao, Q.: Comparative competitive coding for personal identification by using finger vein and finger dorsal texture fusion. Inf. Sci. **268**, 20–32 (2014)
17. Mian, A.: Illumination invariant recognition and 3D reconstruction of faces using desktop optics. Opt. Express **19**, 7491–7506 (2011)

Steganography and Cryptography for User Data in Calendars

Péter Vörös, Péter Hudoba and Attila Kiss

Abstract Privacy is the ability of an individual or group to seclude themselves, or their private information, and thereby express themselves selectively. Nowadays the increasing usage and popularity of web services indicate the likelihood of privacy disclosures. Data breaches are the number one privacy threat when using cloud services. To avoid the possibility for data theft users should not trust anyone including the service providers. If they do not rely on the providers' defense methods, they have to make sure that outgoing traffic remains a secret for everyone. In this paper, we propose our model and our prototype implementation for a client-side data encoder, that aims to solve the privacy disclosure problems. We show different binarization (character string to binary string) and encoding methods (binary string to character string), which with, we can provide secure and hidden data encoding by translating the user data into different sentences that can be considered as valid data. We make sure that only encoded data can be stored in the providers' storage, therefore we not just ensure that no one has direct access to unencoded data, but we are able to conceal the very existence of the encoding at all.

Keywords Privacy · Web service security · Public cloud · Personal data management · Data breach

1 Introduction

Users generate a high amount of personal data daily by using online services, such as social networks, search engines, etc., without asking themselves the question of how valuable is my data for the providers? Because of the fact that collecting and

P. Vörös (✉) · P. Hudoba · A. Kiss
Eötvös Loránd University, Budapest, Hungary
e-mail: vopraai@inf.elte.hu

P. Hudoba
e-mail: peter.hudoba@inf.elte.hu

A. Kiss
e-mail: kiss@inf.elte.hu

M. Huk et al. (eds.), *Intelligent Information and Database Systems: Recent Developments*, Studies in Computational Intelligence 830,
https://doi.org/10.1007/978-3-030-14132-5_19

classifying personal data is a crucial part of targeted advertising, the answer is: very valuable. Due to the marketers increasing reliance on customer data, its value has significantly grown in the last decade. More and more data are being stored in various cloud storage and they remove the burden of personally taking care of its security since most of these service providers do not offer any options for them to set up trusted encryption. In this paper, we aim to show a possible way how to hide personal data from the curious ones.

We have created a two-layer security technique to protect the personal data of the users. The first layer is the steganography, which is the practice of concealing a file, message, image, or video within another file, message, image, or video. We use it to prevent the adversary from even noticing if the information they accessed is fake or real. Katzenbeisser et al. present steganography in details in [1]. We are using a dictionary-based algorithm, creating an encoded text data, that looks like a normal user input to trick the adversary and hide the sensitive data. The second layer is a simple personal password-based encryption that encodes the data with a simple stream cipher.

There are different already existing solutions for common users or smaller companies, but there are two major problems with those: lack of trust and the need to get used to a new application. In our prototype, the user can use their ordinary application with an extension, without the need for new software. Our prototype runs invisibly in the background, while users are able to access the service the same way as they can do without the extra layers of security.

This paper is structured as follows: In Sect. 2. we analyze the Related work of this topic. We show our model and different approaches in Sect. 3., and a performance analysis between these different methods in Sect. 4. Later we show our prototype in Sect. 5. Finally in Sect. 6. we conclude our work, and in Sect. 7. we collect the possible future improvements that can be made.

2 Related Work

Security plays a significant role in all Web applications, especially in the cloud era, when providers offer a wide range of useful services to users. Privacy issues in public clouds had huge attention throughout the past years. It is well known and researched in several papers, that cloud services introduce several new possible ways of attack. Shaikh et al. in [2] collect many areas of security challenges in the cloud. They identified that security is the biggest hurdle in wide acceptance of cloud computing such as Data loss, Leakage of Data, Client's trust, User's Authentication, Malicious users handling, Wrong usage of Cloud computing and its services, Hijacking of sessions while accessing data.

However, there number of security-aware advanced users who are able to realize the potential vulnerabilities and privacy issues in web applications is extremely low due to several publications such as [3, 4]. In our previous work, we analyzed some of the biggest threats: the Account Hijacking [5], and DoS/DDoS problem [6, 7]. In this paper we set the focus on Data Breaches [8] which is stated as the highest threat among those.

Definition 1 A Data Breach is an incident in which sensitive, protected or confidential information is released, viewed, stolen or used by an individual who is not authorized to do so [9].

There are several different ways of securing private data, one possible classification can be server-, and client-side protection.

2.1 Server-Side Protection

By server-side protection, we mean that the user trusts the destination and sends its data without any additional modification (e.g.: encryption) and the server stores it somehow securely.

Pearson et al. in [10] proposed a privacy manager that relies on obfuscation techniques. The key in this application is to store encrypted data on the server-side only. With this privacy manager, one can reduce the amount of sensitive information stored in untrusted servers. However, this technique may only be a working solution if the providers would accept to use it.

Song et al. in [11] introduced a Data Protection as a Service solution which is a suite of security primitives offered by a cloud platform.

In [12] Antonenkov et al. define methods and systems of encrypting files at a client in cloud base environments.

Due to the fact that providers' systems are like a black box to end-users, our problem is that no one can be 100% sure about what cloud providers' state as inaccessible to others is really secure.

2.2 Client-Side Protection

By implementing an extra layer of security over existing services users will have full control of their data. Client-side protection makes sure that no raw or usable data goes to the 3rd party servers.

Chow et al. [13] proposed FHE (Fully Homomorphic Encryption) to keep the privacy of the data in cloud systems. The usage of FHE enables computation on encrypted data, that is stored in the cloud provider's servers. This encrypted data can be processed without decryption, while the cloud providers will have no knowledge about the input data, the processing function, the result, and any intermediate result values.

In [14] Sanamrad et al. propose a very similar aspect as we do, instead of the client-side protection they use a security middleware to encrypt all contents of the outgoing traffic to Google Calendar and Gmail, by this they are able to store no unencrypted data in Google cloud. Security middleware makes the process transparent to users, however, they need to set up the middleware by themselves to make it work.

ShadowCrypt is proposed by He et al. in [15] is a client side as a browser extension, we use the same approach. However, their goal was to store encrypted data in cloud servers but our is to produce a human-readable encrypted output. By doing so we are able to hide the fact that the data is encrypted.

3 Our Framework

Our encoding process has three steps, first, we convert all of the input data into binary strings, encrypt with a simple stream cipher, and then convert the binary data to a user readable information again. In this section, we will describe formally the process and discuss functions for the steps. Let us denote with $\mathbb{K} = \{a, b, \dots z, \_\}$ the set of alphabetic characters and whitespace. The _ sign denotes the whitespace character. Let $\mathbb{B} = \{0, 1\}$ the set of possible values of one "bit". Denote the concatenation operator with $||$ (e.g. "a" $||$ "b" = "ab" and "1"$||$"0" = "10"). The input of the algorithms can be defined as a $\mathbb{K}^*$, that contains finite many alphabetic/whitespace characters concatenated. We will refer to a $\mathbb{K}^*$ as string.

We have an infinite sequence of $binarize_i : \mathbb{K}^* \to \mathbb{B}^* \times \mathbb{N}$ ($\forall i \in \mathbb{N}^+$) functions, those converts a string text into a binary sequence and gives back how many characters has been used from the beginning of the string.

Denote our input as $t_1 t_2 t_3 \dots$, where $t_i \in \mathbb{K}$. In the process step of the algorithm we calculate after each other the $binarize$ functions and shift left the input string with the number of consumed characters. For example the result of this process can be $binarize_1(t_1||t_2||\dots)||binarize_2(t_3||t_4||\dots)||binarize_3(t_4||t_5||\dots)\dots$, where the first function used two, the second one characters.

The second step is encryption to avoid information claim for the adversary if the steganography failed. This step we use simply a stream cipher. Denote this cipher with $encrypt$.

The third step of one process is encoding the binary string into readable information again. Use an infinite sequence of $encode_i : \mathbb{B}^* \to \mathbb{K}^* \times \mathbb{N}$ ($\forall i \in \mathbb{N}^+$) functions as a decoding functions and do the same in a inverted way. The functions consume bits from the binary string after each other the $encode$ functions and the end of the step we get a character string.

In this paper, if we refer to a binary string ($\mathbb{B}^*$) as a natural number, that means we simply convert the binary string to a natural number (e.g. as a big-endian unsigned integer) and vica-versa. In formulas we will use $binaryToNumber$ and $numberToBinary$ functions.

On Algorithm 1 we can see the top-level overview of the algorithm (The Algorithm 2 shows a subprocedure of the main function).

Algorithm 1: Overview of the encoding

Function `MAIN(`*input, key, seed*`)`**:**
- **Data**: $input$ - the input string that has to be encoded
- **Data**: key - the encryption key
- **Data**: $seed$ - the encryption seed
- **Result**: $result$ - the encoded string
- $temp \leftarrow ENCODE(input, (binarize_i)_{i \in \mathbb{N}^+})$
- $encrypted \leftarrow encrypt(temp, seed, key)$
- $result \leftarrow ENCODE(encrypted, (encode_i)_{i \in \mathbb{N}^+})$

Algorithm 2: One step in the modell

Function `ENCODE(`$input, (fun_i)_{i \in \mathbb{N}^+}$`)`**:**
- **Data**: $input$ - the input string that has to be encoded
- **Data**: $(fun_i)_{i \in \mathbb{N}^+}$ - sequence of functions for encoding
- **Result**: $result$ - the encoded string
- $temp \leftarrow input$
- $i \leftarrow 1$
- $result \leftarrow$“”
- **while** *temp is not empty* **do**
  - $binstring, n \leftarrow fun_i(temp)$
  - shift left $temp$ string with n characters
  - $result \leftarrow result || binstring$
  - $i \leftarrow i + 1$

3.1 Binarization Functions

Trivial binarization The trivial way to binarize the text is using a character table (e.g. ASCII, UTF, or custom) and return the binary value of that character as a binary string. In practice we use a custom character table that maps the $\mathbb{K}^*$ directly to $\mathbb{N}$ (_ $= 0$, $a = 1$, $b = 2, \ldots$).

Huffman compressor The $binarize$ function can be a Huffman compressor, that has a fixed compression tree by an alphabet frequency and always consumes 1 character. We used the English alphabet frequency from Table 1 of [16].

Dictionary based Because of our steganography layer, we have to have some dictionaries, so we can use them to try to compress the input data in the first step. As preparation, we create a union of our dictionaries and sort by length. The *binarize* function consume one word (separated by whitespace character) at once and searches for that word in the union of the used dictionaries. If a concrete word has been found in the dictionary, it uses the index instead of the word if the result is shorter than the trivial binarization of the word. We have to denote that in the string uniquely if we are using an index, so when we find an improvement, we add a special character and the index encoded into a string. This means if a word is not in the dictionary or not causes improvement this approach falls back to the trivial binarization method for that word, so the result cannot be worst than trivial binarization unless the custom character table is the power of 2.

Remark 1 The dictionary and Huffman compressor will not be used together, because we try to take advantage of the distribution of the characters and the dictionary-based binarization goes to uniform distribution.

3.2 Encryption

For encryption step, we use simple stream cipher and we use the calendar item id as the seed to avoid the similar encryption of a calendar message. To protect the data we ask a key from the user that will be used for the stream cipher.

3.3 Encoding Functions

Base64 The trivial way to transform the binary string into user readable information that using base64 encoding, but this approach does not satisfy the aim of this paper, therefore, the existence of the encoding is clearly visible.

Dictionary based Every *encode* function has its own dictionary. For all of the functions we have a (not necessarily disjoint) set $S \subset \mathbb{K}^*$ dictionary and we consume $\lfloor log_2(|S|) \rfloor$ bits from the binary string. By the natural number of the binary string i, we give back the dictionary element with index i.

We used different words for verbs, nouns, adjectives, names, and places to achieve realistic sentences.[1] If we denote the encode functions based on these dictionaries

[1] Sources: https://www.talkenglish.com/vocabulary/top-500-adjectives.aspx (Reached: 2018-09-14) https://www.talkenglish.com/vocabulary/top-1000-verbs.aspx (Reached: 2018-09-14) https://www.talkenglish.com/vocabulary/top-1500-nouns.aspx (Reached: 2018-09-14) https://github.com/dominictarr/random-name/blob/master/names.txt (Reached: 2018-09-14) https://github.com/dominictarr/random-name/blob/master/places.txt (Reached: 2018-09-14).

with $encode_{verbs}$, $encode_{nouns}, \ldots$ and $encode_{comma} : x \mapsto$ (",", 0), we can use the following function sequence:

$$enc_i = \begin{cases} encode_{verbs} & i \equiv 0 \ (mod\ 4) \\ encode_{adjectives} & i \equiv 1 \ (mod\ 4) \\ encode_{places} & i \equiv 2 \ (mod\ 4) \\ encode_{comma} & i \equiv 3 \ (mod\ 4) \end{cases}$$

Branched dictionary based There are typical calendar titles for common users that can hardener our data hiding. To achieve this we created branching. We have a sequence of encode function sequences $(encode_i)_j \ j \in 1, \ldots, m$ and depends on the first (of the input part where the branching starts) $\lfloor log_2(m) \rfloor$ which sequence will be used to that binary string. One possible branch that we can use is:

$$Meet/Call \xrightarrow{\text{with}} NAME \xrightarrow[\text{togetherwith}]{\text{and/or/along}} NAME \xrightarrow[\text{to}]{\text{for}} NOUN \rightarrow VERB \xrightarrow[\text{at/on}]{\text{in/by}} PLACE$$

An example output of this branch is: Meet with Anna and Joe for administration reduce at Thompsonville.

4 Comparison

In this section, we will demonstrate the effectiveness of the different combinations of the binarization and encoding methods. It is irrelevant to compare to the existing cloud privacy solutions, because the novelty in our approach is to combine steganography techniques with cloud security and to our best knowledge this approach has not got any candidate.

If we consider combinations, we can state that the Huffmann binarizing method has to be at least good as the trivial one in average if the code tree good for the character distribution of the source strings. We take advantage of Huffman coding by assigning shorter bit sequences for characters with smaller occurrence probability.

If we use the Dictionary method as binarizing, we also will have shorter bit sequence with higher probability. We substitute a word with an id if it is shorter than the original, but the range of the function has to be bigger, therefore it can be encoded in a longer bit sequence only. Nonetheless in most of the cases it is better than a trivial binarizing.

The effectiveness of Dictionary + Dictionary method can be seen in Table 1. In the second row, we show the Base64 result just to demonstrate a non-human-readable output which is an easily identifiable encoding. Due to this, the comparison between the Base64 results and the Dictionary based results do not really make sense. Because of the word "meeting" is an element of our dictionaries, it can be compressed and

Table 1 The results and efficiency of the different encodings on "meeting". This is a special case when the original message appears in one of our dictionaries, therefore Dictionary + Dictionary method achieves a lot better result than the others

Method	Result	Size
Original	Meeting	
Trivial + Trivial	5QKj4k7	
Trivial + Dictionary	Go away with Jemie to web then	7
Huffman + Dictionary	Go away to Gratia for nature round	7
Dictionary + Dictionary	Dare annual Aaronsburg	3

Table 2 The results and efficiency of the different encodings on "go to mall". Due to the bigger complexity of the sentence there is no significant difference between the algorithms

Method	Result	Size
Original	Go to mall	
Trivial + Trivial	2Ey57WQ5Z0	
Trivial + Dictionary	Call with Mimi togetherwith Jackie to ladder short at	9
Huffman + Dictionary	Melli mission about safe with Rhody by Kailua	8
Dictionary + Dictionary	Call with Mimi togetherwith Jackie to ladder short at	9

encoded to only 3 words which is less than 45% of the required space we needed to represent the result of the Huffman compression. We can say that for short one word texts (which is really common as a calendar event) Dictionary + Dictionary can achieve very nice results.

As one can see on Table 2. the original sentence "go to mall" is 10 characters long. Dictionary methods result a more or less "valid sentence" and while they satisfy the same requirements the number of words needed to represent their results can be a good metric for comparison.

The different binarization methods highly affect the length of the outputs. As can be seen in Table 3. Huffman compression gives a better result than the trivial method in most cases. It's because of the fix compression tree, with the English alphabet frequency, that we use in our Huffman compressor. If a sentence contains mostly "rare characters" (e.g.: j, q, u, z) the result can be worse than the trivial (e.g.: Table 3. "schuss" row).

In most of the cases the Huffman + Dictionary works really well with a non-specific calendar text based code tree already.

Table 3 Comparing algorithms effeciency on different sentences

Binarize	Trivial	Huffman	Dictionary
Encode	Dictionary	Dictionary	Dictionary
Call dan	8	6	7
See tom	7	7	6
Do work	7	6	7
Meeting	7	7	3
Go to mall	10	8	9
Italian course	13	10	14
Lunch with elena	15	13	19
Lunch with a very important person	30	40	42
Meet with anna and bruce	21	25	32
Dinner with anna	15	12	14
cccccc	6	7	6
Hello world	11	8	11
auf wiedersehen	14	12	14
Schuss	6	7	4
Transfer secret data to John	24	27	32
Call + 1234 56789912	6	7	6
Wipe hard drives	15	14	15
Meeting with dr. peter	19	19	21
Poker party	11	9	8

5 Prototype

For our prototype implementation, we used Yahoo calendar as the target service. Our client-side encoding script is written in JS and injected into the HTML page by Greasemonkey, a browser extension. By the choice of Greasemonkey, the portability is not an issue, because both major browsers support either Greasemonkey or Tampermonkey which is the same program called on a different name.

The users can not tell the difference between the original website, and the one with our JS code injected, because our script runs completely in the background. As one can see on Fig. 1 creating new entries, modification and deletion work the same as without the usage of our script. No functionality is lost due to the extra layer of security.

The script listens for each submit, modification event and extracts the user input from the fields, encodes it (with any of the methods stated above) and then lets the original process to finish. This will ensure that no unencrypted data is uploaded

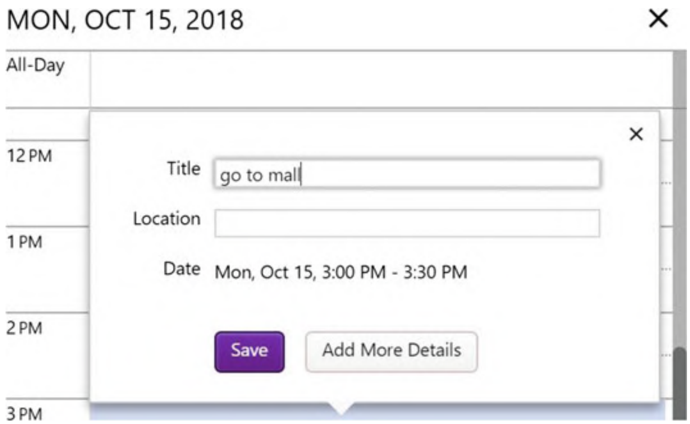


Fig. 1 Creation of a new entry while using our script does not differ from the original Yahoo calendar usage

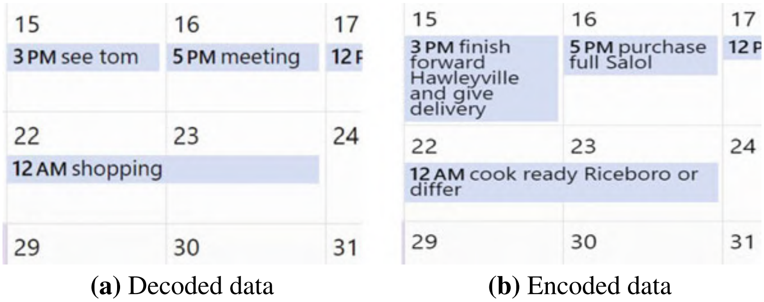


Fig. 2 This fig shows how the calendar looks like if one sees it while using the prototype, and on the left side is what the provider and the potential attackers may have access to

to the cloud servers. Similar things happen when the page is reloaded, or some modification comes from the server. The script decrypts the message that was stored on the servers. This is how we can keep private user data hidden from the servers, and how we manage not to cause any inconvenience to users.

The difference between the original or decoded and the encoded data can be seen on Fig. 2.

6 Conclusion

In this paper, our goal was to show a possible way to defend against data theft and make providers unable to have direct access to our personal data, by using steganography to be able to hide the encoding itself. Our proposed model uses 3 steps: binarization, encryption, encoding to produce a very effective way of hiding private data due to the fact that it can provide a secure and hidden data encoding, by translating the user data into different sentences that can be considered as valid data. We compared the efficiency of the stated approaches and showed how these can be integrated into a prototype.

7 Future Work

People make mistakes. These typos, for example an extra space before a comma. In the future, we would like to use this great source of information to claim new bits in the system.

We plan to improve our binarize algorithms to fill the gaps caused by the non 2 power sized domain of the functions. If we have a function that domain does not a power of 2, we do not use some bit sequences. If we enumerate all of the combinations together of the domains we can claim a better algorithm.

We would also like to collect a calendar sentence database, to make more valid values as encoded value and get a better input for dictionary-based binarization.

We will use various semantical structure to encode the binary data.

Acknowledgements Supported by the ÚNKP-18-3 New National Excellence Program of the Ministry of Human Capacities.
The project was supported by the European Union, co-financed by the European Social Fund (EFOP-3.6.3-VEKOP-16-2017-00002).

References

1. Katzenbeisser, S., Petitcolas, F.: Information Hiding Techniques for Steganography and Digital Watermarking. Artech house, Boston (2000)
2. Shaikh, F.B., Haider, S.: Security threats in cloud computing. In: 2011 International Conference for Internet Technology and Secured Transactions (ICITST), pp. 214–219. IEEE (2011)
3. Mylonas, A., Kastania, A., Gritzalis, D.: Delegate the smartphone user? Security awareness in smartphone platforms. Comput. Secur. **34**, 47–66 (2013)
4. Huang, D.-L., Rau, P.-L.P., Salvendy, G.: Perception of information security. Behav. Inf. Technol. **29**(3), 221–232 (2010)
5. Vörös, P., Kiss, A.: Tookie: a new way to secure sessions. In: Recent Developments in Intelligent Information and Database Systems, pp. 195–207. Springer, Berlin (2016)
6. Csubák, D., Szücs, K., Vörös, P., Kiss, A.: Big data testbed for network attack detection. Acta Polytech. Hung. **13**(2), 47–57 (2016)
7. Vörös, P., Laki, S., Kiss, A.: Distributed firewall on dataplane against ddos attack. 4th Winter School of Ph.D. Students in Informatics and Mathematics, p. 37
8. The treacherous twelve cloud - computing top threats in 2016 (2016). https://cloudsecurityalliance.org/group/top-threats
9. U. D. O. HEALTH, H. S. A. for Children, and Families, "State and tribal child welfare information systems, information security data breach response plans. Technical report (2015)

10. Pearson, S., Shen, Y., Mowbray, M.: A privacy manager for cloud computing. In: IEEE International Conference on Cloud Computing, pp. 90–106. Springer, Berlin (2009)
11. Song, D., Shi, E., Fischer, I., Shankar, U.: Cloud data protection for the masses. Computer **45**(1), 39–45 (2012)
12. Antonenkov, L., Romanovskiy, S., Uraltsev, N., Prokofiev, A.: Client-side encryption, US Patent App. 14/271,918, 18 Dec 2014
13. Chow, R., Golle, P., Jakobsson, M., Shi, E., Staddon, J., Masuoka, R., Molina, J.: Controlling data in the cloud: outsourcing computation without outsourcing control. In: Proceedings of the 2009 ACM Workshop on Cloud Computing security, pp. 85–90. ACM, New York (2009)
14. Sanamrad, T., Nick, P., Widmer, D., Kossmann, D., Braun, L.: My private google calendar and gmail. IEEE Data Eng. Bull. **35**(4), 83–92 (2012)
15. He, W., Akhawe, D., Jain, S., Shi, E., Song, D.: Shadowcrypt: encrypted web applications for everyone. In: Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security, pp. 1028–1039. ACM, New York (2014)
16. Jones, M.N., Mewhort, D.J.: Case-sensitive letter and bigram frequency counts from large-scale english corpora. Behav. Res. Methods Instrum. Comput. **36**(3), 388–396 (2004)

Part IV
Intelligent Systems in Biomedicine

Modeling and Extraction of Retinal Blood Vessels from RetCam 3 Based on Morphological Segmentation

Alice Krestanova, Jan Kubicek, Juraj Timkovic, Marek Penhaker, David Oczka and Jan Vanus

Abstract This paper deals with the analysis and modeling of the retinal blood vessels system. The aim of the analysis is the design and implementation of a fully automated segmentation model based on the morphological segmentation, allowing for extraction of the blood system area within the binary model, where other retinal structures are suppressed. An important feature of the model is sensitivity and robustness to declare the efficacy of segmentation in an environment with worse image parameters. For this reason, the designed model is also tested for data where the vascular system is visualized under a low contrast. Part of the analysis is the comparative testing of the designed model against selected segmentation methods based on objective criteria. The designed model was tested and verified on dataset from system RetCam 3 containing 22 images.

Keywords Retina · Morphological segmentation · Detection · Blood vessels · Retinopathy of prematurity · Comparison · RetCam 3 · MATLAB

A. Krestanova · J. Kubicek (✉) · M. Penhaker · D. Oczka · J. Vanus
FEECS, VSB-Technical University of Ostrava, K450, 17. Listopadu 15, Poruba, Ostrava, Czech Republic
e-mail: jan.kubicek@vsb.cz

A. Krestanova
e-mail: alice.krestanova@vsb.cz

M. Penhaker
e-mail: marek.penhaker@vsb.cz

D. Oczka
e-mail: david.oczka@vsb.cz

J. Vanus
e-mail: jan.vanus@vsb.cz

J. Timkovic
Clinic of Ophthalmology, University Hospital Ostrava, Ostrava, Czech Republic
e-mail: juraj.timkovic@fno.cz

M. Huk et al. (eds.), *Intelligent Information and Database Systems: Recent Developments*, Studies in Computational Intelligence 830,
https://doi.org/10.1007/978-3-030-14132-5_20

1 Introduction

Retinopathy of prematurity (ROP) represents a vasoproliferative disease, which affects the immature retinal blood vessels of the retina in prematurity infants. This disease occurs in children, who belong into the following risk groups: children born before the 32nd gestation week, children with a birth weight of less than 1.500 g. Another possible risk of disease with this disease is the hypoxemia or hypercapnia of the newborn infants. For the ROP diagnostics, RetCam 3 is conventionally used. It is a device from Clarity Medical Systems that allows for retrieve the retina of premature infants. Images from RetCam 3 have a resolution 640 × 480 pixels. At the images, there are observable objects as the retinal blood vessels, optical nerve and retinal lesions. One of the symptoms of this disease is the tortuosity or blood vessel winding. Up to now, the tortuosity has been evaluated by the ophthalmologists by manually or by visualizing multiple images taken by the RetCam 3 instrument. For this reason, it is necessary to automatically extract the vascular system from image with a view to further evaluation of tortuosity. Since the tortuosity was not represent by a parameter, which it corresponds with the blood vessels curving. It is need to extract the blood vessels in a binary segmentation model. On the binary model of blood vessels is applied an extraction of the curvature parameters would allowing for estimation and classification the tortuosity [1–3].

2 Related Work

This paper deals with image processing of retinal images taken by the RetCam 3. In related work have been processed predominantly images taken by fundus camera with higher resolution (605 × 700 pixels vs. RetCam 3 640 × 480 pixels) [4, 5].

Therefore, the recently reported segmentation methods do not have to be effectively applied in the RetCam 3 images [6, 7].

This fact predetermines that low resolution images represent a challenging area for the medical image segmentation [8]. Method uses extraction of the green channel from color image was taken by RetCam 3, noise filter is applied and mapping vascular system of retina on black background [9].

In different approach an image has been pre-processed by grayscale image conversion, Gaussian blur, Laplace edge detection, noise removal, image binarization and removal of isolated areas. Segmentation has been enhanced by branch detection techniques and endpoint detection [10].

Segmentation of retinal blood vessels based on Gabor's transformation. In this preprocessing image, extraction of the green chrominance component of image, image sharpening, histogram adaptation and median filtration were separated. Gabor's transformation utilized the frequency and spatial approximation of the retinal vessel system. Gabor's filters use Gabor's wave, which reliably detects the shape of the blood vessels system [11].

Segmentation methods which have been applicate on images taken by fundus camera are described below. Segmentation of retinal blood vessels on the base of method clustering K-means [12]. Segmentation based on tracking is applied to local operators representing the blood vessels system. Blood vessel tracking begins at the initialization point, detects the center of the blood vessel lines or their surroundings by analyzing the orthogonal pixels in the tracking directions [13].

Supervised method has been used training sets. The original image is processed by wavelet operation that allows to divide image into several levels based on the extraction features. The tracking method is based on combination of vessel directions and eight Hessian matrix vectors used for segmentation of the vascular system [14, 15]. Segmentation of the retinal blood vessels using neural networks utilizing a 7D features vector composed of gray shades and feature representing the invariant moment. The multilayer forward neural network is used for training and classification [5].

Unsupervised method can be used using a fuzzy C-means clustering algorithm, that uses language descriptions to decide whether it is or it is not a vessel. Fuzzy tracking is based on the determination of membership of functions in two language values [16, 17].

Different approach image pre-processing, when the RGB image was converted to the HSI color model. Additionally, it was used an adaptive histogram to increase the contrast and average filter to remove a noise. The next step for segmentation image is Top-hat transformation [18].

3 Proposal of Morphological Segmentation of Retinal Blood Vessels

This segmentation method is based on a sequence of the morphological operations. It has been tested on 22 retinal images of prematurity infants with the retinopathy of prematurity taken by the RetCam 3. The retinal images are divided into 11 contrast images and 11 non-contrast images. We have deliberately performed testing on worse quality images in order to evaluate objective effectivity and robustness of the morphological image segmentation.

In the first step of the algorithm for segmentation of retinal blood vessels is morphological preprocessing, when vascular system is segmented from other anatomical retinal structures in images (e.g. optical disc). The aim of the morphological preprocessing is emphasized vasculature and preserved transitions of the bifurcation of the vascular system. The algorithm has been used difference between supreme I_c and infim I_B of morphological opening operation applied to the original image with two linear structural elements with different sizes [19].

An image I_c can be obtained by the detected supreme from the original image I_0 with linear structural elements with many different rotations, where B is meaning structural element in rotation i:

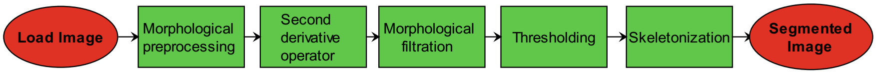


Fig. 1 Algorithm for modelling and extraction of retinal vascular system based on morphological segmentation

$$I_c = sup_{i=1,\ldots,12}\{\gamma_{B_i}(I_0)\} \tag{1}$$

Removing linear shapes from the image corresponds to replacing them (local) backgrounds. The local background of the image is obtained by opening the boundary of the linear length B by a length in many ways, where i is represented 12 structural elements in 0–165°.

$$I_B = inf_{i=1,\ldots,12}\{\gamma_{B_i}(I_0)\} \tag{2}$$

Filtered image I_v, which is contained only linear shapes:

$$I_V = I_C - I_B \tag{3}$$

The result of this morphological preprocessing is the emphasis on linear structure (blood vessels). Second derivative operator is used to emphasizing thin vascular structures. Morphological filtration is removed nonlinear bright parts, which are not belong to retinal blood vessels, where G_θ is represented 1D Gaussian convolution kernel with angle θ and $G''_{\theta+90}$ is 1D second derivation of Gaussian convolution kernel in angle $(90 + \theta)^\circ$ (see Eq. 4).

$$I_l = R_{I_{diff}}\left(sup_{i=1,\ldots,12}\left\{\gamma_{B_i}\left(sup_{\theta=0\ldots180}\left(I_v * G_\theta * \left(G''_{\theta+90}\right)\right)\right)\right\}\right) \tag{4}$$

Then binary mask of image is created by the hysteresis thresholding. Two binary images I_{low} and I_{high} are created: I_{low} is created by thresholding with low thresholding value t_{low} and I_{high} is created by thresholding with high thresholding value t_{high}. The segmentation output is represented by the binary segmentation mask, where the white pixels represent the detected blood vessels, contrarily the adjacent retinal structures are suppressed, as black. Skeletonization of the mask allows to identify points in the image, where blood vessels are crossed (bifurcations, transitions). [19] The entire image segmentation process is shown below (see Fig. 1).

The setting parameters were mostly the same, only setting of the threshold value was changed depending on the contrast of the image (Table 1). In some cases, the length of the linear smoothing filter was changed to achieve maximally sensitize detection of vascular system at minimal noise levels. Figure 2 shows a segmentation of the retinal blood vessels from native image, when setting two different threshold values. Images show that when lower threshold values are set, the image also shown fine background vessels, while with higher threshold setting, the image shows the major vessels of the vascular system and eliminates the smaller blood vessels (Fig. 3).

Table 1 Parameters settings for the segmentation procedure of the blood vessels system

Parameters	Values
Linear length of the structural element for vascular amplification	11
Linear length of the structural element for the removal of blood vessels	11
Length of linear smoothing filter	7–15
Threshold value of tagged image (higher value)	15–50
Threshold value of mask image (lower value)	10–40
Defining removing objects that are smaller than (in pixels)	180
Fill holes smaller than (in pixels)	10

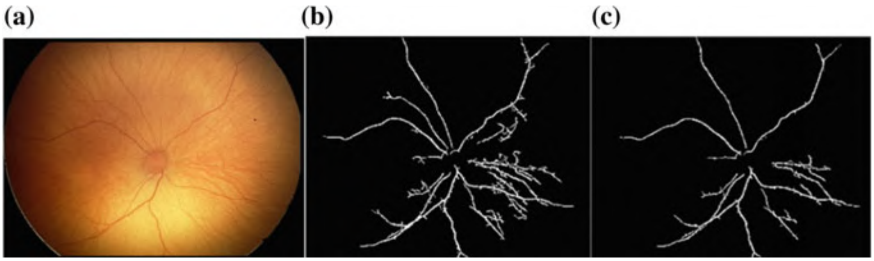


Fig. 2 Native image (**a**), segmented image with setting threshold parameters (15, 25) (**b**), segmented image with setting threshold parameters (30, 50) (**c**)

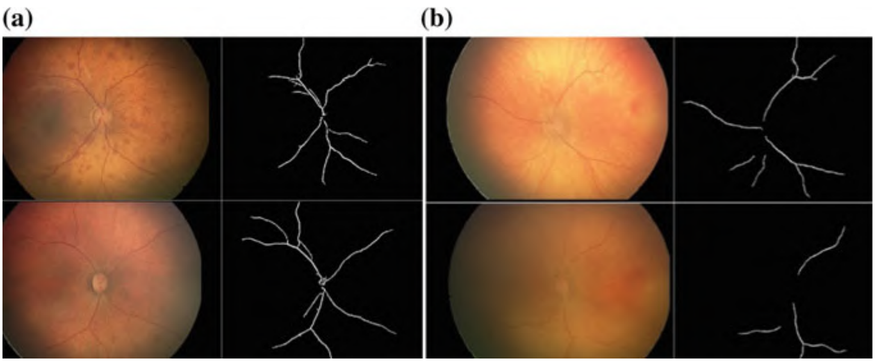


Fig. 3 Comparative segmentation analysis of contrast retinal images (**a**) and non-contrast images (**b**)

4 Quantitative Comparison and Testing

They have been compared segmentation methods based on: morphological operations, the Gabor transformation, the Top-hat and Bot-hat transformation in contrast with the proposed segmentation method in this paper. The output image *x* after the tested segmentation and output image *y* after segmentation using proposed morphological operation based on the Heneghen method are objectively compared by using parameters for objective evaluation of the segmentation methods. In the comparative analysis, we are using both contrast and non-contrast RetCam 3 images in order to find out robustness and effectivity of the proposed segmentation method.

4.1 Definition of Objectification Parameters

A mean squared error, structural similarity index and image correlation are used for the objective comparison of applied image segmentation methods.

Structural similarity index (*SSIM*) is based on computing brightness, contrast and structure. Index is multiplicative combination of these parameters. *SSIM* takes values in the scale from 0 to 1, while the higher value means greater the similarity between images *x* and *y*.

Mean squared error (***MSE***) is calculated between images *x* and *y*. This parameter is used to evaluate the accuracy of a measurement, in this case a segmentation. It is true, that the lower the MSE value, the smaller disagreement between images.

Image correlation indicates a size of the linear correlation between images *x* and *y*. It is true, that the higher the correlation coefficient, the higher is correlation. That means images are more linearly similar.

4.2 Results of Quantitative Testing

The Table 2 summarizes the averaged results of the objective evaluation of the proposed method against alternative segmentations. Based on the summary comparison, it can be deduced that, morphological operation and Bottom-hat segmentation are closest to the reference method based on parameters values below. From the values of the Gabor transformation, there is a high image mismatch in contrast to the proposed method. These results may be due to the high sensitivity of the wavelets, which leads to highlighting the noise and its presence in the segmentation model.

The advantage of the proposed segmentation algorithm is that, there is not necessary specific preprocessing of the retinal image. When using the preprocessing methods, the native clinical images may be improved regarding their quality. On the other hand, each preprocessing step modifies the native clinical information which is partially suppressed. Also, the advantage of this method is the absence of noise in a segmented model of vascular system. In this way the morphological segmentation appears to be sensitive, and robust against the image noise.

Table 2 Comparison of alternative segmentation methods against proposed segmentation method

Segmentation method	SSIM	MSE	Correlation coefficient
Morphological operation	0.88	0.03	0.13
Gabor transformation	2.60×10^{-5}	3.05×10^{4}	0.01
Bottom-hat segmentation	0.86	2.19×10^{3}	0.44

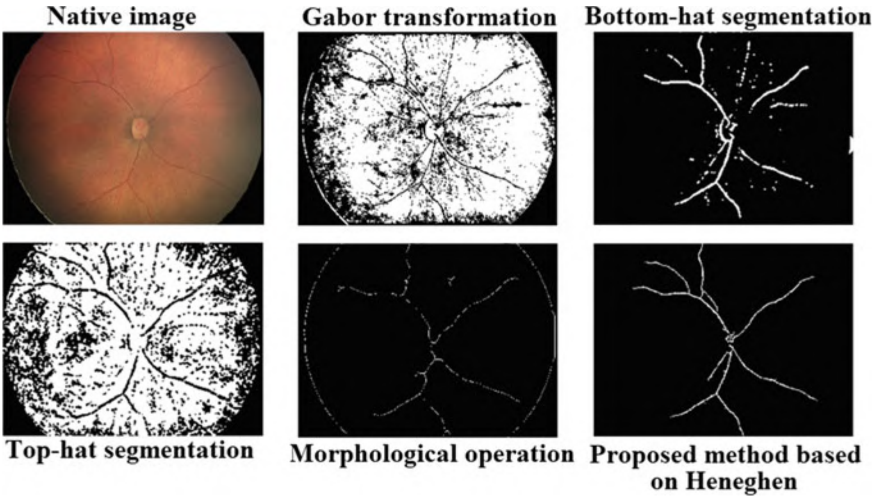


Fig. 4 Comparative analysis of alternative segmentation techniques applied to contrast image in contrast to proposed method

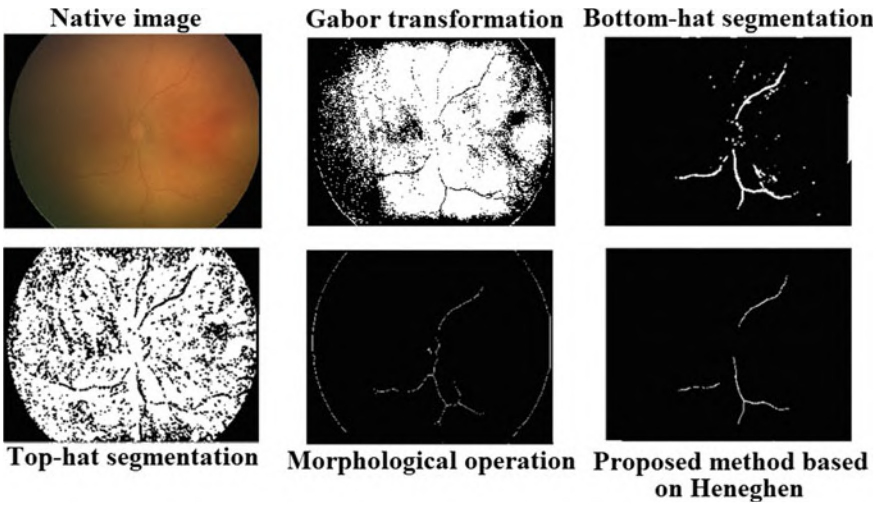


Fig. 5 Comparative analysis of alternative segmentation techniques applied to non-contrast image in contrast to proposed method

The disadvantages of used retinal images are their high variability, which caused by a different contrast between the blood vessels and the background, unequal illumination in the scanned images by a wide field and high visibility of the choroidal vessels associated with insufficient pigmentation of the infants choroidal artery. Another problem is a low resolution of images 640×480 pixels due to a wide-screen image. Also, movements artifacts can be created in images due movement of infants eye. Figures 4 and 5 represent a comparison amongst the proposed segmentation model and alternative segmentation models.

5 Conclusion

In the clinical ophthalmology, the blood vessels segmentation belongs to one of the most frequent tasks being necessary regarding a proper diagnosis of the ROP. This disease is clinically examined by the RetCam 3, which generates data in lower resolutions. Such fact predetermines weaker contrast amongst individual retinal structures including the retinal blood vessels.

Based on the aforementioned reasons, segmentation method for such retinal records should be sensitive, and in the same time robust against the image noise and artefacts. The analyzed method based on the morphological segmentation is able to reliably identify individual blood vessels even in the noisy environment. In order to justify that, we have analyzed contrast and non-contrast retinal images. When comparing the segmentation results with alternative segmentation methods, the morphological segmentation method suppresses the most from the image noise. Quality of the segmentation can also be measured on the base of the contour of the retina. When comparing the proposed method with others, the retinal contour is suppressed. This method effectively separates blood vessel's tree, while other retinal structures are suppressed.

In the future time, we are going to focus on the segmentation model features extraction. The blood vessels model should serve as a classification tool allowing for the tortuosity differentiation. We suppose calculation a set of the curvature features, which will be calculated from the proposed model. Such feature vector will serve as a base for classification and prediction of the pathological signs of the tortuosity.

Acknowledgements The work and the contributions were supported by the project SV4508811/2101Biomedical Engineering Systems XIV'. This study was also supported by the research project The Czech Science Foundation (GACR) 2017 No. 17-03037S Investment evaluation of medical device development run at the Faculty of Informatics and Management, University of Hradec Kralove, Czech Republic. This study was supported by the research project The Czech Science Foundation (TACR) ETA No. TL01000302 Medical Devices development as an effective investment for public and private entities. This article was supported by the Ministry of Education of the Czech Republic (Project No. SP2018/170). This work was supported by the European Regional Development Fund in the Research Centre of Advanced Mechatronic Systems project, project number CZ.02.1.01/0.0/0.0/16_019/0000867 within the Operational Programme Research, Development and Education.

References

1. Kubicek, J., Timkovic, J., Krestanova, A., Augustynek, M., Penhaker, M., Bryjova, I.: Morphological segmentation of retinal blood vessels and consequent tortuosity extraction. J. Telecommun. Electron. Comput. Eng. **10**(1–4), 73–77 (2018)
2. Jomier, J., Wallace, D.K., Aylward, S.R.: Quantification of retinopathy of prematurity via vessel segmentation. In: Ellis, R.E., Peters, T.M. (eds.) Medical Image Computing and Computer-Assisted Intervention—MICCAI 2003, LNCS, vol. 2879. Springer, Berlin (2003)
3. Chee, R., Patel, S.N., Jonas, K.E., Gupta, M.P., Cambell, J.P., Chiang, M.F., Chan, R.V.P.: Current trends in telemedicine for retinopathy of prematurity. J. Ophthalmol. **16**(1) (2017)
4. Jan, J., Odstrcilik, J., Gazarek, J., Kolar, R.: Retinal image analysis aimed at blood vessel tree segmentation and early detection of neural layer deterioration. Comput. Med. Imaging Graph. **36**(6), 431–441 (2012)
5. Fraz, M.M., Remagnino, P., et al.: Blood vessel segmentation methodologies in retinal images—a survey. Comput. Methods Programs Biomed. **108**(1), 407–433 (2012)
6. Li, M., Ma, Z., Liu, C., Zhang, G., Han, Z.: Robust retinal blood vessel segmentation based on reinforcement local descriptions. BioMed Res. Int. **2017**, Article ID 2028946, 9 pages (2017)
7. Patwari, M.B., Manza, R.R., Rajput, Y.M., Saswade, M., Deshpande, N.: Automatic detection of retinal venous beading and tortuosity by using image processing techniques. Int. J. Comput. Appl. (0975 – 8887) Recent Advances in Information Technology (2014)
8. Poletti, E., Florin, D., Grisan, E., Ruggeri, A.: Automatic vessel segmentation in wide-field retina images of infants with retinopathy of prematurity. In: 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Boston, MA, USA (2011)
9. Jayadev, C., Vinekar, A., Mohanachandra, P., Desai, S., Suveer, A., Mangalesh, S., Bauer, N., Shetty, B.: Enhancing image characteristics of retinal images of aggressive posterior retinopathy of prematurity using a novel software (RetiView). BioMed Res. Int. (2015)
10. Turior, R., Chutinantvarodom, P., Uyyanonvara, B.: Automatic tortuosity classification using machine learning approach. Appl. Mech. Mater. **241–244**, 3143–3147 (2013)
11. Kubicek, J., Kosturikova, J., Penhaker, M., Augustynek, M., Kuca, K.: Segmentation based on Gabor transformation with machine learning: modeling of retinal blood vessels system from RetCam images and tortuosity extraction. In: New Trends in Intelligent Software Methodologies, Tools and Techniques, vol. 297, pp. 270–283 (2017)
12. Mapayi, T., Tapamo, J.R., Vipiri, S., Adio, A.O.: Automatic retinal vessel detection and tortuosity measurement. Image Anal. Stereology **35**(2), 117–135 (2016)
13. Kaur, M., Talwar, R.: Review on: blood vessel extraction and eye retinopathy detection. Int. J. Comput. Sci. Inf. Technol. **5**(6), 7513–7516 (2014)
14. Gregory, S., Trevor, D., Piotr, I.: Nearest-Neighbor Methods in Learning and Vision: Theory and Practice (Neural Information Processing). MIT Press, Cambridge (2006)
15. Xu, L., Luo, S.: A novel method for blood vessel detection from retinal images. BioMed. Eng. Online **9**, 14 (2010)
16. Tolias, Y.A., Panas, S.M.: A fuzzy vessel tracking algorithm for retinal images based on fuzzy clustering. IEEE Trans. Med. Imaging **17**, 263–273 (1998)
17. Kande, G.B., Subbaiah, P.V., Savithri, T.S.: Unsupervised fuzzy based vessel segmentation in pathological digital fundus images. J. Med. Syst. **34**, 849–858 (2009)
18. Nazari, A., Mustafa, M., Zulkifley, M.: Segmentation of retinal blood vessels by top hat multiscale detection for optic disc removal. J. Teknologi. **77**(6), 47–53 (2015)
19. Heneghan, C., Flynn, J., O'Keefe, M., Cahill, M.: Characterization of changes in blood vessel width and tortuosity in retinopathy of prematurity using image analysis. Med. Image Anal. **6**(4), 407–429 (2002)

Modelling Medical Information and Knowledge with OWL and Topic Maps

Martina Husáková

Abstract Majority of facts about autoimmune diseases are shattered in various web sources. It is difficult to provide fundamental facts about specific autoimmune disease without spending a lot of time. The paper investigates the OWL and the Topic Maps standard for building of an information and knowledge repository including autoimmune diseases. This repository should facilitate findability of facts about autoimmune diseases. The OWL and the Topic Maps is compared with the RDF and the RDFS model for answering question which approach is more suitable for development of the repository.

Keywords Autoimmunity · Navigation · Semantics · OWL · Topic Maps · RDF(S)

1 Introduction

Difficulties in investigation of an immune system lie in its complexity that is caused by an existence of huge amount of players and interrelations between them. These players (immune cells and immune organs) take part in innate and adaptive immune processes for ensuring homeostasis of a living entity. The immune system needs not behave always correctly. Autoimmune diseases (ADs) are possible consequences of this inadequate behaviour when the immune system damages own cells and tissues. There exist more than 80 types of ADs and many of them share similar signs [1].

Our research was focused on the exploration of ADs with the multi-agent systems. We consulted our conceptual [2] and computational models [3] with an immunologist and used medical literature and web sources for development of these models. We found out that if a non-professional wants to receive well-arranged, easily navigable, compact and understandable pieces of information about specific autoimmune dis-

M. Husáková (✉)
Faculty of Informatics and Management, University of Hradec Králové, Hradec Králové, Czech Republic
e-mail: martina.husakova.2@uhk.cz

M. Huk et al. (eds.), *Intelligent Information and Database Systems: Recent Developments*, Studies in Computational Intelligence 830,
https://doi.org/10.1007/978-3-030-14132-5_21

ease (AD), he (she) has to visit a lot of web sources for obtaining sufficient amount of information. These facts are not often interlinked with other facts. Only lists of keywords with short descriptions are often provided. Various forms of information presentation are often missing (e.g. images, videos, synonyms or audio files).

This paper continues in a research that has been presented in [4] where the RDFS (Resource Description Framework Schema) ontology of ADs is introduced. This paper investigates the OWL (Ontology Web Language) and the Topic Maps standard for building repository of ADs and compares them with the RDF (Resource Description Framework) and the RDFS for evaluation of their similarities, differences and usability for development of the repository of ADs.

The paper is structured as follows. Section 2 introduces various sources where the facts about ADs can be found on the web. Section 2 also presents the requirements for the repository of ADs. Section 3 introduces developmental process of the OWL-based and the Topic Maps-based repository. Section 4 includes comparison of the RDF, the RDFS, the OWL and the Topic Maps. Section 5 discusses proposed solution together with future directions and Sect. 6 concludes the paper.

2 Web Sources About Autoimmune Diseases

There is a huge amount of information sources about autoimmune diseases on the web. If we would like to find information about specific autoimmune disease, we can meet the following possibilities in the web space most often:

- the lists of medical keywords with short descriptions (e.g. Harvard Medical Dictionary [5] or eMedicineHealth.com vocabulary [6]),
- the lists of medical keywords with deeper explanation of these keywords (Medical Encyclopedia at MedlinePlus [7]),
- the information pages with long descriptions (e.g. LabTestsOnline.org [8]).

Computational biomedical ontologies are specific category used for formal representation of biomedical terms. These graph-based structures are shared as vocabularies between people, people and machines and between machines. These structures facilitates communication and problem solving. The following biomedical ontologies includes facts about various diseases including ADs: *Human Disease Ontology* (DOID) [9], *Experimental Factor Ontology* (EFO) [10], *Medical dictionary for Regulatory Activities* (MEDDRA) [11], *International Statistical Classification of Diseases and Related Health Problems* (ICD) [12], *Human phenotype ontology* (HP) [13] or *MeSH* thesaurus [14].

It is find out that insufficient navigation between information sources and mutual interconnectedness of these sources about ADs is rare on the web. Machine-processable information (knowledge) and representation of semantics of medical concepts should be offered for provision of relevant answers to user queries. It is found out that no less or more formal structure that is presented above satisfies requirements which are presented below. These requirements are based on intensive study of previously mentioned sources presenting facts about autoimmune diseases:

- Content-related criteria:
  - general description of an autoimmune disease,
  - symptoms of an autoimmune disease,
  - possible cause(s) of an autoimmune disease,
  - examination of an autoimmune disease,
  - possible treatment (medication) of an autoimmune disease,
  - prevention against an autoimmune disease,
  - relations to other autoimmune diseases,
  - synonyms of an autoimmune disease (if exist),
  - additional information about an autoimmune disease for its deeper understanding (web sources: interesting web pages, pdf files, video links, links to audio sequences),
  - target audience: non-expert users.
- Design and implementation-related criteria:
  - user-friendly navigation between various facts,
  - machine-readable information/knowledge about autoimmune diseases,
  - contextual knowledge representation,
  - representation of semantics for machine understanding and processing of medical concepts.

These requirements became the main criteria for development of the information and knowledge repository including specific ADs. The following section presents modelling approaches which are considered for development of the repository.

3 Development of Ontology-Based Repositories

3.1 OWL-Based Repository and Implementation Issues

The first repository is based on mixture of the RDF and the RDFS [4]. The second version (AutoimmuneDiseases.owl) extends the RDFS ontology about concepts used for semantics modelling. This ontology represents ADs in the OWL 2 [15]. The OWL is an official W3C standard extending the RDF [16] and the RDFS model [17]. Thanks to the formal logic, the OWL (2) can be used for modelling semantics of concepts more precisely in comparison to the RDF and the RDFS, see Sect. 4 and Table 3. Figure 1 depicts fundamental building blocks used for development of the OWL 2

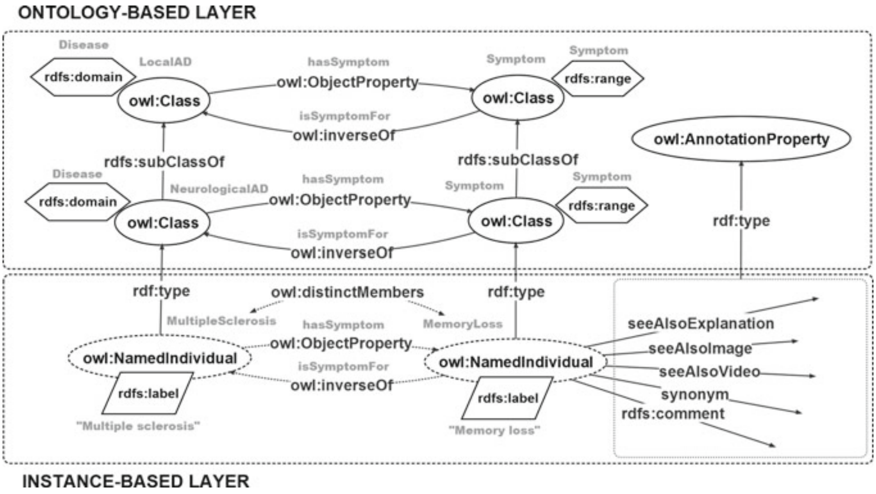


Fig. 1 Structure of the OWL 2 ontology of autoimmune diseases

ontology. The examples are depicted in a grey colour for differentiation. The figure is simplified for better interpretation.

The OWL structure consists of the following core building blocks:

- *OWL class* is a collection of specific instances sharing some similarities.
  - Examples: AutoimmuneDisease, Medication, Prevention, Cause, Characteristics, Symptom, Test (how an AD can be diagnosed).
- *OWL individual* is a concrete instance of a class.
  - Examples: MultipleSclerosis (OWL class: NeurologicalAD), Ciclosporin (OWL Class: ImmunosuppressiveDrug), CT scanning (OWL class: Test).
- *OWL property* represents relations between OWL classes (OWL individuals). Domains and ranges of these properties represent between which classes (individuals) these properties can exist.
  - Examples: (domain: Disease) hasCause (range: Cause), (domain: Disease) hasSymptom (range: Symptom), occurs with (domain, range: Disease).
- *OWL annotation property* is a specific relation used for adding additional information (metadata) into the ontology.
  - Examples: a comment (a description of a medical term), a synonym, a web source—a seeAlsoExplanation, a seeAlsoImage and a seeAlsoVideo, a label (a human readable version of a medical term).

The OWL 2 ontology is developed in the Protégé 5.2.0—one of the most used editor for ontological engineering. The OntoGraf plugin is used for visualisation of

Table 1 Statistics of the RDFS and the OWL 2 ontology

Ontology metric	Count in	
	the RDFS ontology [4]	the OWL 2 ontology
Ontological classes	30	37
Properties	8	15
Annotation properties	–	8
Reflexive properties	–	1
Individuals	229	229
Property assertions	353	402
Annotation assertions	–	373
Definable classes	–	5

the ontological structure. The OWL 2 ontology is consistent, i.e. no logical contradictions are detected by a debugger plugin that is available in the Protégé 5.2.0. Statistics of the RDFS and the OWL 2 ontology AutoimmuneDiseases.owl is the following, see Table 1.

3.2 Topic Maps-Based Repository and Implementation Issues

The third version of the repository is developed with the Topic Maps. Topic Maps are the ISO standard (ISO/IEC 13250) used in the information and knowledge management with the main aim to improve navigation and findability of relevant information and knowledge on the web [18]. The standard proposes "infrastructure" consisting of the knowledge-based layer and the information-based layer, see Fig. 2. The figure is simplified for better interpretation. The knowledge-based layer (ontology) contains knowledge which is represented by "building stones" of each Topic Map—a document playing a role of a metadata layer existing above various digital sources. The information layer consists of information sources linked to the elements in the knowledge layer. Following "building stones" are distinguished by the standard and used in the proposed repository (AutoimmuneDiseasesTM.ltm):

- *Topic* is a digital representation of any "thing" that is in the centre of our interest.
  - Examples: Uveitis, MultipleSclerosis, SchirmerTest, Corticoid, Ibuprofen.
- *Association* represents relationship among topics.
  - Examples: hasSymptom, isCausedBy, isDiagnosedBy, isTreatedBy.
- *Occurrence* connects an information source (e.g. a web page) with a specific topic. An external occurrence connects externally available information source with specific topic, e.g. a web page (an external source) describing AD. An internal

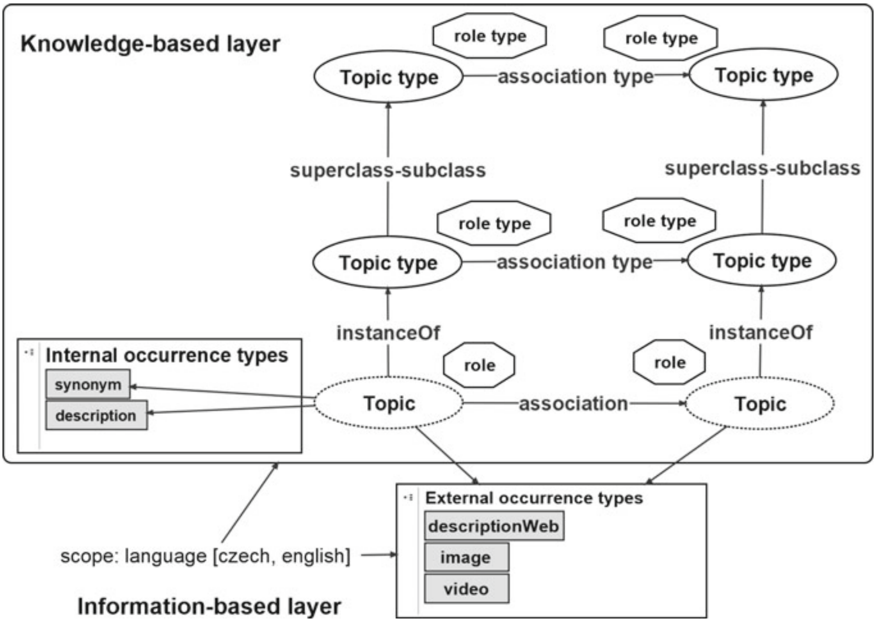


Fig. 2 Structure of the Topic Map of autoimmune diseases

occurrence is available only from a Topic Map, e.g. a synonym for a word or a brief description of AD.

These Topic Maps elements have types. Types (classes) represent containers into which particular topics, associations and occurrences below according to their characteristics, e.g.:

- *Topic type*: Disease, HematologicAD, Drug, Cause, Symptom.
- *Association type*: superclass-subclass relation used for modelling hierarchy of topic types.
- *Internal occurrence type*: a synonym or a description.
- *External occurrence type*: an image (a hyperlink for an image), a description on the web (descriptive information about specific medical term on the web), video (a hyperlink for a video sequence).

The Topic Map is developed in the non-standardised syntax LTM (The Linear Topic Map) [19]. This notation permits to represent facts by simple structures without losing meanings. The LTM-based Topic Map is visualised in the Omnigator—a component of the open-source tool Ontopia [20]. Statistics of the Topic Map is the following, see Table 2.

Table 2 Statistics of the Topic Map AutoimmuneDiseasesTM.ltm

Element of the Topic Map	Count
Topic types	18
Association types	7
Internal occurrence types	3
External occurrence types	3
Topics	238
Associations	379
Internal occurrences	53
External occurrences	125

4 RDF, RDFS, OWL and Topic Maps—Comparison

4.1 Purpose

The RDF is mainly used for representation of metadata for web sources. It models statements about different sources without their structure. The RDFS extends the RDF. It is used for representation of collection of resources (classes) and for building of the RDF vocabularies (ontologies). It provides structure for our data without their semantics. The OWL (2) adds the semantics into the classes and the RDF(S) statements. On the basis of the description logics, it can be used for inferring new facts in an ontology and checking its consistency. All of these approaches are used for making web content more accessible to machines where intelligent agents can use these structures for problem solving, decision making and communication. On the other hand, the Topic Maps are more designed for humans, i.e. for improvement findability of facts and navigation between them.

4.2 Modelling Concepts

The Topic Maps specification uses an element named topic which is a machine-processable representation of a concept. It can represent whatever, i.e. a digital or a non-digital source. Topics are represented as resources in the RDF(S). The OWL uses individuals for modelling instances of classes (topics). It can be seen that compared formalisms use different terminology for instances of classes. Topics are grouped according similarities into the topic types (classes). Classes are also used by the RDF, the RDFS and the OWL, but the RDF is not able to model their hierarchy.

The Topic Maps standard uses associations for modelling relations between topics, topics and occurrences. Associations can be represented in the RDF, the RDFS and the OWL, but specific distinctions exist between them. Associations are named as properties in these formalisms. The RDF provides "built-in" rdf:type property for modelling instances of classes. Hierarchy of classes and properties can be represented only in the RDFS, the OWL and the Topic Maps.

The RDFS introduces a domain and a range concept. A domain states that any individual that has a given property is an instance of one or more classes specified in a domain. A range states that the values of a property are instances of one or more classes specified in a range. These concepts are not built in the Topic Maps model, but concept of roles is very similar to them. A role models context of an association, i.e. it represents participation of specific concepts in an association. As an example, "psoriasis can be treated by phototherapy". Psoriasis is a role player of a role dermatologic disease and phototherapy is a role player of a role medication.

The context is also represented by a scope concept that is defined only in the Topic Maps. The scope is a constraint in which something is said about specific topic, e.g. the scope can be used for filtering of fundamental concepts (for beginners) and more advanced concepts (for experts). The scope is not directly supported by the RDF(S) and the OWL, but the OWL uses annotation property adding metadata to concepts. As the example for filtering concepts according to a difficulty: "a hasTargetAudience" annotation property with two possible values "novice" and "expert" can be added into the ontology. Software application can query concepts of the ontology suitable only for the experts.

Occurrences represent additional information about concepts in the Topic Maps. External occurrences can be represented as resources in the RDF(S) and the OWL, e.g. rdfs:seeAlso property represents further information about the subject resource. Internal occurrences can be represented as values of rdfs:comment property in the RDFS (the OWL). External and internal occurrences can be represented as the values of an annotation properties in the OWL. Occurrence types are classes of the information sources. These can be modelled as the classes in the RDFS and the OWL.

Association types (classes) represent nature of a relation and participation of a topic in the association. Association classes are not typically used only in the Topic Maps, but in the UML (Unified Modelling) language. They are not defined by the RDFS or the OWL specification. n-ary design pattern is used for this purpose. This design pattern is also applied in reification.

Reification is directly defined by the Topic Maps. A topic is utilized as a subject that can be connected with a non-topic subject (e.g. an association). It is possible to model the n-ary relationships in a Topic Map or to add metadata for a Topic Map as a whole. Reification is not directly supported by the RDF, the RDFS or the OWL, but blank nodes can be used for this purpose.

Level of semantics representation is one of the fundamental differences between presented approaches. The OWL is the most expressive language. It provides representation of disjointness. The OWL looks more into "the inner structure" of the classes. Additional conditions (with restrictions or logical constructors) can be added into the classes for specification of their semantics. As the example, the OWL 2 ontology of ADs contains the DermatologicDisease class. What is the meaning of this class? Two formal conditions are necessary and sufficient for modelling semantics of this class: *Disease AND influences some Derma*. Meaning is the following: Der-

Table 3 RDF, RDFS, OWL and Topic Maps—comparison

	RDF	RDFS	OWL 2.0	Topic Maps
General comparison				
Year of publication	2004 (W3C Recommendation)	2004 (W3C Recommendation)	2004 (W3C Recommendation, OWL 1.0)	2006 (ISO Topic Maps Data Model)
Purpose	Metadata representation about web resources Data model for data exchange	Representation of simple RDF vocabularies on the web Representation of hierarchy of resources	Representation semantics of data Inference Checking of an ontology consistency	Provision of an intelligent navigation Personalised view on data Information filtration and merging
Target group	Machine (e.g. an intelligent agent, a search engine)	Machine (e.g. an intelligent agent, a search engine)	Machine (e.g. an intelligent agent, a search engine)	Human
Basic building blocks	RDF triples: subject, predicate, object	RDF + class, subClassOf, subPropertyOf, domain, range, datatype, label, comment, seeAlso	RDF and RDFS + disjointness, logic constructors (and, or, not), property (transitive, functional, symmetric, asymmetric, reflexive, irreflexive), restrictions, sameAs, differentFrom	Topic, association, occurrence (external, internal) and their types, role and its type, scope
Level of semantics expression	Low-level	Low-level	High-level	Higher than RDF (RDFS), but lower than OWL
Utilisation for inference	Yes	Yes	Yes	It is possible, but it is not the main purpose
Optimised for	Metadata representation	Inference	Inference	Findability of information
Basic syntaxes	Turtle, RDF/XML, N3, TriG, JSON-LD, N-Quads, N-Triples, RDFa	RDF/XML, Turtle	Turtle, Functional, RDF/XML, Manchester, OWL/XML	XTM, LTM, AsTMa, CTM

(continued)

Table 3 (continued)

	RDF	RDFS	OWL 2.0	Topic Maps
Basic query languages	SPARQL 1.0 (1.1)	SPARQL 1.0 (1.1)	SPARQL 1.0 (1.1) with modifications	tolog
The most cited tools	RDF Studio Protégé TopBraid Composer	RDF Studio Protégé TopBraid Composer	Protégé TopBraid Composer	Ontopia ZTM Wandora
Comparison—modelling view				
Structure	Graph	Graph	Graph	Graph
Associations	Oriented	Oriented	Oriented	Without orientation (roles are used for orientation)
Reification	Not formally defined	Not formally defined	Not formally defined	Defined
Identification	URI	URI	URI	Subject identifier subject locator
Variant names	–	rdfs:label	rdfs:label	Base name (full name), name for sorting, name for displaying (shorter name)
Context representation	Not formally defined	Not formally defined	Not formally defined	Defined (scope)

matologic disease is a disease which influences derma and can influence something else. This statement is opened for next possibilities. Table 3 presents comparison of the above mentioned approaches.

5 Discussion

The RDF, the RDFS, the OWL and the Topic Maps are used for development of the repository of ADs. Why are these approaches used in a development process? Section 3 presents design-related and implementation-related criteria which should be taken into account for repository building. Machine-readability of medical concepts, intuitive navigation, semantics of medical terms and contextual knowledge representation are fundamental requirements for the repository. The HTML 5-based web pages can also be used for presentation facts about ADs, but semantics extensions of the HTML 5 [21] are not sufficient for provision of more expressive semantics of medical statements. A relational database can also be used as a repository of ADs, but a graph-based structure and support of semantics representation cannot be realised with this type of database. The RDF data model can be used for building of a graph database, but semantics of data cannot be precisely defined with this solution. The intended approach that could be somewhere between the Topic Maps and the OWL would be the best choice for development of the repository of ADs where high language expressivity and support for intelligent navigation between web sources could be offered.

6 Conclusion

Conceptual models help to capture fundamental concepts and relations between them in an abstract form. Concepts correspond with abstract or real entities which are in a center of our interest and which can be used in computational models.

Level of formality decides whether a computational system (e.g. an intelligent agent) will be able to use these models, e.g. for decision making, communication, inference or problem solving. The paper is basically focused on two formal approaches for information and knowledge representation. The OWL language (extending the RDF and the RDFS) and the Topic Maps model is applied for building prototype of the repository including facts about autoimmune diseases. The OWL-based prototype is based on the RDFS ontology presented in [4]. These repositories should facilitate findability of information (knowledge) about particular autoimmune disease in one place without necessity to explore various kinds of additional resources which are shattered in the web space. Presented approaches are mutually compared, because it was necessary to find out similarities and differences for a decision which approach is more suitable for development of the repository. The OWL is more formal in comparison to the Topic Maps approach, but it is not directly used for web sites development. Topic Maps are richer in a view of semantics representation in comparison to the RDF (the RDFS) and can be directly used for web sites development.

Future development is going to be focused on how the OWL could be used for web site development which would integrate facts about ADs and reason with them. The Owlready2 is a Python-based library mainly used for management of the OWL structures [22]. Flask is a python-based micro-framework used for web development [23]. These frameworks are going to be used for building of the application. The repositories include 17 autoimmune diseases. This amount is not sufficient. The ontology is going to be extended and integrated into the web application.

Acknowledgements Czech Science Foundation project 18-01246S is kindly acknowledged.

References

1. Delves, P.J., et al.: Roitt's Essential Immunology, 12 edn. Wiley-Blackwell, 560p (2011). ISBN 978-1405196833
2. Husáková, M.: Conceptual Modelling in Computational Immunology, 1st edn. Tomáš Bruckner Publishing (Academic Series), Řepín, Živonín (2015)
3. Cimler, R., Husáková, M., Koláčková, M.: Exploration of autoimmune diseases using multi-agent systems. In: Nguyen, N., Iliadis, L., Manolopoulos, Y., Trawiński, B. (eds.) Computational Collective Intelligence. ICCCI 2016, LNCS Springer-Verlag, vol. 9876, pp. 282–291 (2016)
4. Husáková, M.: Representation of autoimmune diseases with RDFS. In: Nguyen, N., Pimenidis, E., Khan, Z., Trawiński, B. (eds.) Computational Collective Intelligence. ICCCI 2018. LNCS Springer-Verlag, vol. 11055, pp. 43–52 (2018)
5. Harvard Medical School. Medical Dictionary of Health Terms.: Retrieved 20 Apr 2018 from https://www.health.harvard.edu/a-through-c
6. eMedicineHealth.com. Online Medical Dictionary. Retrieved 20 Apr 2018 from https://www.emedicinehealth.com/medical-dictionary-definitions/article_em.htm
7. U. S. National Library of Medicine. MedlinePlus, Medical Encyclopedia. Retrieved 20 Apr 2018 from https://medlineplus.gov/encyclopedia.html
8. Lab Tests Online (AACC).: Retrieved 4 Sept 2018 from https://labtestsonline.org/conditions/autoimmune-diseases
9. Schriml, L.M., et al.: Disease ontology: a backbone for disease semantic integration. Nucl. Acids Res. **40**(Database issue), D940–D946 (2012). http://doi.org/10.1093/nar/gkr972. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3245088/
10. Malone, J., et al.: Modeling sample variables with an experimental factor ontology. Bioinformatics **26**(8), 1112–1118 (2010). http://doi.org/10.1093/bioinformatics/btq099. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2853691/
11. The International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use.: Understanding MedDRA: The Medical Dictionary for Regulatory Activities (2013). Retrieved 10 Mar 2018 from https://www.meddra.org/sites/default/files/page/documents/meddra2013.pdf
12. World Health Organization.: ICD-10: International Statistical Classification of Diseases and Related Health Problems (10th rev.), 2 (2011). Retrieved 12 Mar 2018 from http://www.who.int/classifications/icd/ICD10Volume2_en_2010.pdf
13. Köhler, S., et al.: The human phenotype ontology project: linking molecular biology and disease through phenotype data. Nucl. Acids Res. **42**(Database issue), D966–D974 (2014). http://doi.org/10.1093/nar/gkt1026

14. U.S. National Library of Medicine: Fact Sheet: Medical Subject Headings (MeSH). https://www.nlm.nih.gov/pubs/factsheets/mesh.html (cit. 3. 9. 2018)
15. W3C. OWL 2 Web Ontology Language: Document Overview (Second Edition), W3C Recommendation 11 December 2012. Retrieved 25 Apr 2018 from https://www.w3.org/TR/2012/REC-owl2-overview-20121211/#
16. W3C. RDF 1.1 Concepts and Abstract Syntax, W3C Recommendation 25 February 2014. Retrieved 25 Apr 2018 from https://www.w3.org/TR/2014/REC-rdf11-concepts-20140225/
17. W3C. RDF Schema 1.1, W3C Recommendation 25 February 2014. Retrieved 26 Apr 2018 from https://www.w3.org/TR/2014/REC-rdf-schema-20140225/
18. ISO: ISO 13250-2 Topic Maps—Data Model. Retrieved 5 Apr 2018 from http://www.isotopicmaps.org/sam/
19. Garshol, L.M.: The Linear Topic Map Notation. Definition and introduction, ver. 1.1 (2001). Retrieved 12 Mar 2018 from http://www.regnet.org/members/demo/ontopia/doc/misc/ltm.html
20. Ontopia. Ontopia homepage. Retrieved 26 Apr 2018 from http://svn.mssm.nl/
21. W3C. HTML 5.2, W3C Recommendation, 14 December 2017. Retrieved 23 Apr 2018 from https://www.w3.org/TR/html5/
22. Lamy, J.B.: Owlready: ontology-oriented programming in Python with automatic classification and high level constructs for biomedical ontologies. Artif. Intell. Med. **80**, 11–28 (2017)
23. Flask. Retrieved 21 Dec 2018 from http://flask.pocoo.org/

Comet Assay Classification for Buccal Mucosa's DNA Damage Measurement with Super Tiny Dataset Using Transfer Learning

Afiahayati, Edgar Anarossi, Ryna Dwi Yanuaryska, Fajar Ulin Nuha and Sri Mulyana

Abstract Comet assay or single cell gel electrophoresis assay (SCGE) is a method which is frequently used to measure the damage of DNA. The results of comet assay is a set of comet images, then the comet images are classified to measure the level of DNA damage. Currently, there are several softwares for comet assay image analysis, both free and commercial. Commercial software is quite expensive, while free software is limited, especially for buccal mucosa cell and super tiny comet image dataset. In this research, we propose a classification model for comet assay with super tiny image dataset which is taken from buccal mucosa cells. We propose a transfer learning based convolutional neural network (CNN) model. We have compared the transfer learning model with CNN-support vector machine (SVM) and ordinary CNN. In our experiments, we use super tiny dataset consisting of 73 images. Our transfer learning model gives an accuracy 70.5%, while CNN-SVM gives 62.3% and ordinary CNN gives 63.5%. We also compare our transfer learning model with most frequently used, free comet assay analysis software, OpenComet. Open-Comet gives an accuracy 11.5%. Our transfer learning model is promising for comet assay for buccal mucosa cell and super tiny dataset.

Keywords Transfer learning · Tiny dataset · Comet assay · Buccal mucosa cell

1 Introduction

Comet assay is an excellent method for detecting DNA damage due to its versatility and sensitivity [1]. Comets can be identified and scored by visual inspection by an expert, but this method is very time-consuming and subjective. Therefore, image

Afiahayati (✉) · E. Anarossi · F. U. Nuha · S. Mulyana
Department of Computer Science and Electronics, Faculty of Mathematics and Natural Sciences, Universitas Gadjah Mada, Yogyakarta, Indonesia
e-mail: afia@ugm.ac.id

R. D. Yanuaryska
Department of Radiology Dentomaxillofacial, Faculty of Dentistry, Universitas Gadjah Mada, Yogyakarta, Indonesia

M. Huk et al. (eds.), *Intelligent Information and Database Systems: Recent Developments*, Studies in Computational Intelligence 830,
https://doi.org/10.1007/978-3-030-14132-5_22

analysis software packages were developed [2]. Several tools have been developed to analyze comet assay images. Most of the tools use geometric measuring of the comet's core and tail [3–5]. Currently, there are some softwares for comet assay image analysis, both free and commercial. Commercial software is quite expensive, while free software is limited, especially for buccal mucosa cell. Comet image from different sample sources, for instance mucosa swab and cell line, have slightly different geometric algorithm. It affects the accuracy of the data. Comet images data are hard to obtain, therefore the dataset is limited.

In this research, we propose, an artificial intelligence method specifically developed to score the comet image obtained from buccal mucosa cells. There are two main challenges for this problem domain: (1) comet images obtained from buccal mucosa cells are not like ideal cells since there is a swab activity; (2) the dataset is very tiny. One of artificial intelligence algorithms that could be used to solve image classification problems are Convolutional Neural Network (CNN). CNN is a machine learning method that is designed to process visual information [6–10]. CNN could classify an image based of the patterns learned at the training image sets. The learned patterns then could be used to classify new images. CNN has good performance with a big dataset, while our comet image dataset is super tiny. Therefore, we present a transfer learning based convolutional neural network (CNN) model to classify comet images to measure the level of DNA damage. This work is an extended version of authors' work [11].

2 Data and Methods

2.1 Comet Assay Data

The used comet assay images are cropped image from a microscope slide images of buccal mucosal cells. The total number of images are 73 images. There are 5 levels of the DNA damage. Level 0 represents the lowest damage (the best DNA) while level 4 represents the highest damage (the worst DNA). Those levels are represented as classes in our classification models. Example of comet assay's microscope slide image and samples of comet assay images for each class could be seen in Figs. 1 and 2.

2.2 Pre-processing

In this research, there are 5 stages of pre-processing that produced 2 type of datasets: (1) dataset without data augmentation, and (2) dataset with data augmentation. Pre-processing stages are as follows:

Image Size Normalization Image size normalization uses the Gaussian Pyramid Multi-Scaling Analysis [12] process to expand or reduce the size of image to get as

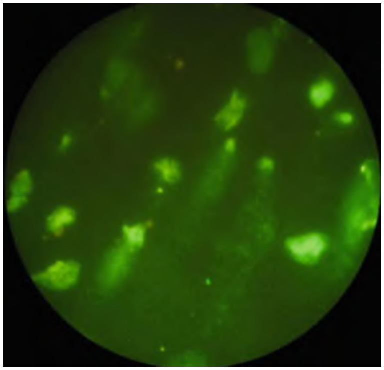


Fig. 1 Comet assay's microscope slide image

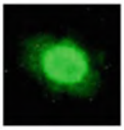
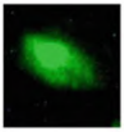
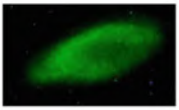
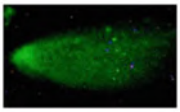
Class	*Comet Assay Image*
0	
1	
2	
3	
4	

Fig. 2 Samples of comet assay images for each class

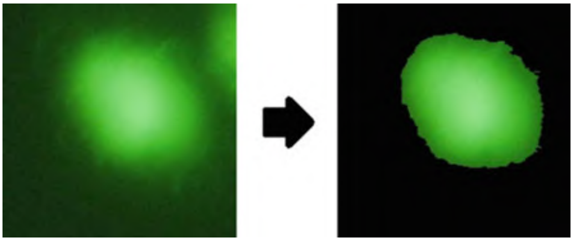


Fig. 3 Comet assay's segmentation

close to 100 × 100 pixel. After that a resize process is done using bilinear interpolation to reach 100 × 100 pixel.

Noise Data Removal Removing images from each class with a similarity percentage to other images in the same class with lower than the predefined similarity threshold. This process is used because the size of dataset are very small, so if the noise images are learned, there will be no other image that could be used to validate if the images are from the respective classes.

Image Segmentation Comet assay objects are separated from the background or other comets inside the image. This process is done to ease the machine learning model at learning the comet assay images. Segmentation is done by using threshold parameter at HSV color space. An example of the segmentation result could be seen at Fig. 3.

RGB Color Space Conversion to Grayscale Color Space Converting the image's RGB color space to Grayscale color space because for the most part of the image, only Green dimension of the RGB has positive values, meanwhile the Red and Blue dimension are zeroes.

Data Augmentation Some of the augmentations used are horizontal and vertical flipping, random cropping, random distortion, and resize. Each classes are augmented to 50 samples per class.

2.3 Classification Models

In order to get good accuracy from such a super dataset, in this research we tried several classification models that could process image data, as follows: (1) Ordinary Convolutional Neural Network, (2) CNN-Support Vector Machine, (3) Transfer Learning (our proposed method).

Ordinary Convolutional Neural Network (CNN) CNN is one of deep learning algorithm variants that is used to process 2 dimensional data. Some parameters that affects a CNN models are convolution layer, pooling layer, activation function, data being used, and how the training is done. Architecture of the CNN model used in this research is represented in Table 1 and Fig. 4.

Table 1 Specification of the CNN model

Layer	Information
Input layer	Image size 100 × 100
Convolution layer	64 filter, size 15 × 15
Pooling layer	Size 2 × 2, stride 2, 2
Convolution layer	256 filter, size 11 × 11
Pooling layer	Size 2 × 2, stride 2, 2
Dense layer	1024 neuron
Dense layer	1024 neuron
Dropout	Dropout rate = 0.2
Output layer	Softmax with 5 neuron

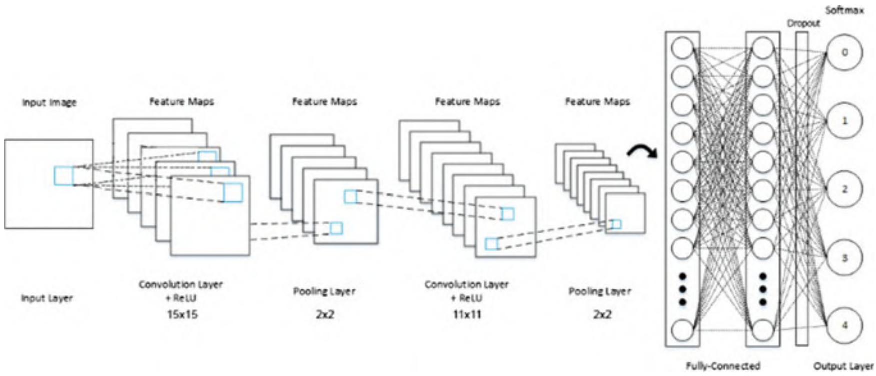


Fig. 4 The architecture of CNN model

The training process is done using the error value calculated from a loss function that is cross entropy and accelerated using ADAGRAD optimizer.

CNN-Support Vector Machine (SVM) SVM solves problems by making N − 1 hyperplane to separate data features with N feature dimensions [13]. Hyperplane or support vector is a line that separates features with as wide margin as possible. Each data will then be categorized into a class according to it's features from the hyperplanes.

Some parameters that affect SVM with RBF as kernel are C and gamma that was tuned in this research. C is a parameter that affects the magnitude of penalty given for each of the misclassified data, meanwhile gamma is a parameter that affects how far the influence of single training example reaches. Best parameters of the SVM model are obtained by using Grid Search. The range of parameters used at grid search is shown in Table 2. The input of SVM in this research is same with the input of the first dense layer at CNN model.

Table 2 Range of parameters used at grid search

Parameter	Range
C	[0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2]
Gamma	[0.0000001, 0.000001, 0.00001, 0.0001, 0.001, 0.01, 0.1, 1]

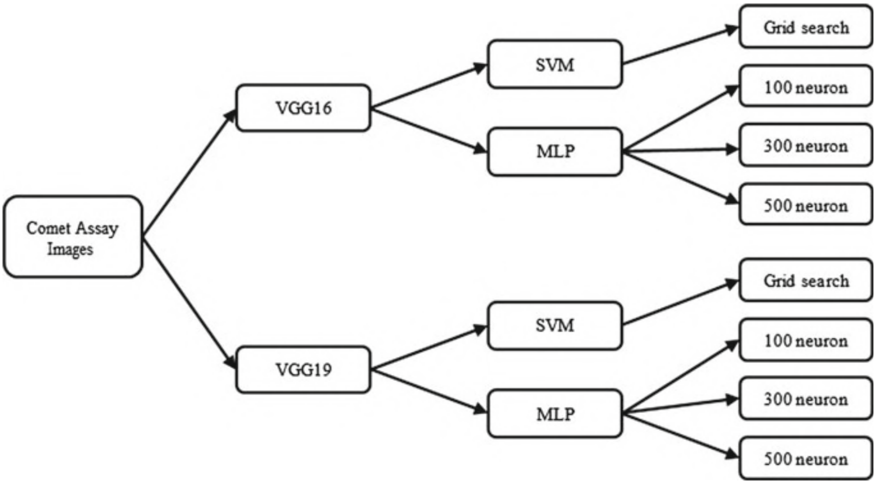


Fig. 5 Classification process using transfer learning

Transfer Learning Transfer learning or inductive transfer is a technique that focused at storing knowledge obtained from solving a problem and applying it to another problem. Transfer learning can be applied at deep learning if a feature learned by a model is general enough to be applied at a target data.

There are two kinds of approach that could be used in transfer learning, that is (1) building a model from scratch and (2) using a pre-trained model. In this research we use a pre-trained model of VGG16 and VGG19 architecture that had been trained on ImageNet dataset which includes 1000 classes [14]. To be able to use the transfer learning model's outputs at classifying comet assay, we used VGG16 and VGG19 outputs as an input at the next classifier that is Multi Layer Perceptron with 1 hidden layer or a Support Vector Machine to get the desired 5 class. Classification process using transfer learning is shown at Fig. 5.

3 Results and Discussion

3.1 Experiment Configuration

For every classification models being tested, 2 types of datasets are used, that is: (1) dataset without data augmentation which includes 61 images, and (2) dataset with data augmentation which includes 250 images. Validations are done using k-fold cross validation with a k value of 4. This research is implemented in Kubuntu 16.10. The CPU is Intel Core i7-6700 CPU @ 3.40 GHz with 16 GB of RAM and only for Ordinary CNN, we use parallel processing with GPU 1080 Ti.

3.2 Results of Ordinary CNN

For the CNN model, testing is exclusively split into 2 parts, which is: (1) testing of parameters related to the quantity of data which includes data augmentation and batch size, and (2) testing of parameters related to the training time which includes learning rate and step size.

The average time of the tests is directly proportional to the size of batch size being tested. Results of data augmentation and batch size tests are shown at Table 3. It is also shown in batch size tests that batch size affects the stability of accuracy and loss function in the training, wherein larger batch size will decrease the amount of change done for each step. From the results obtained from data augmentation and batch size tests at the CNN model, it is shown that unaugmented data has higher accuracy compared to data with augmentations. Furthermore, for the dataset with such a small number, which is $75\% \times 61 = 45$ training data, the best batch size to use in this case is 1 or we could say that the best method for this case is using online learning.

From the results of learning rate and step size experiments, represented at Table 4, it is shown that the best learning rate is 0.00001, whereas using 0.0001 results is a nonoptimal learning while using 0.000001 makes the learning process too slow and

Table 3 Data augmentation and batch size tests results

Batch size	Data	Accuracy	Average time (s)
1	No augmentation	**0.634615406**	237.25
	Augmented	0.500000015	239.5
10	No augmentation	0.519230783	291.25
	Augmented	0.442307696	285
20	No augmentation	0.557692327	291.25
	Augmented	0.519230783	291.25

Table 4 Learning rate and step size tests results

Learning rate	Accuracy (500 steps)	Accuracy (1000 steps)	Accuracy (1500 steps)	Accuracy (2000 steps)
0.0001	0.40385	0.4423	0.40385	0.403846167
0.00001	0.5	0.5577	0.538475	**0.557725**
0.000001	0.365375	0.403825	0.3846	0.403825

Table 5 Best parameter and result of SVM model in CNN-SVM

Data	Accuracy	C	Gamma	Average time (s)
No augmentation	**0.62295082**	1.1	1.00E−07	36
Augmented	0.568	1.5	1.00E−07	345

results in another nonoptimal learning. Too small of a learning rate also results of being stuck in local minima. Size of step size used in this tests isn't too relevant because the step size needed depends on the size of the learning rate.

3.3 Results of CNN-SVM

For the CNN-SVM model, it is shown that most of the tests with unaugmented data results in higher accuracy compared to tests with augmented data. For the gamma parameter, high values result in an optimal model, while small values result in an overfitting model. Furthermore, optimal C value ranges around 0.7–2, so we could say that the penalty given to each misclassification is quite small. The best results of both parameter tests are shown in Table 5.

3.4 Results of Transfer Learning

Transfer Learning model produces the best accuracy compared to the previous models. Similar to previous models, classifications using VGG16 as a feature extractor also results in higher average accuracy when using unaugmented datasets compared to using augmented datasets. From the 4 types of classifiers used to process VGG16's output, it is shown that 1 hidden layer MLP with 300 neuron results in the highest average accuracy. Results of MLP and SVM classification with VGG16 as a feature extractor is shown at Table 6.

Lastly, different from other models' results, Transfer Learning using VGG19 as a feature extractor produces higher accuracy when using augmented datasets rather than unaugmented datasets. This could be caused by 3 extra weight layers that is owned by VGG19 therefore VGG19 could store more features from more data.

Table 6 Results of transfer learning using MLP and SVM as classification methods with VGG16 as a feature extractor

Data/architecture	SVM	MLP 100	MLP 300	MLP 500
No augmentation	0.6202	**0.7049**	**0.7049**	0.639
Augmented	0.6272	0.664	0.68	0.604

Table 7 Results of MLP and SVM classification with VGG19 as a feature extractor

Data/architecture	SVM	MLP 100	MLP 300	MLP 500
No augmentation	0.6238	0.6557	0.623	0.623
Augmented	0.6356	**0.692**	0.668	0.56

Table 8 The accuracy comparison of the overall classification model and OpenComet software [15]

Data/architecture	CNN	CNN-SVM	Transfer learning 1 (VGG16-MLP)	Transfer learning 2 (VGG19-MLP)	OpenComet
No augmentation	0.635	0.623	**0.705**	0.656	0.115
Augmented	0.519	0.568	0.68	0.692	–

From the 4 types of classifiers used to process VGG19's output, it is shown that 1 hidden layer MLP with 100 neuron results in the highest average accuracy. Results of MLP and SVM classification with VGG19 as a feature extractor is shown at Table 7.

The overall model is also compared with the most frequently used, free comet assay analysis software, OpenComet [15]. The data used for OpenComet is same with data for all of our models. The results are represented in Table 8. Our transfer learning 1 model has the highest accuracy compares to other methods. Our transfer learning model gives an accuracy 70.5%, while CNN-SVM gives 62.3% and ordinary CNN gives 63.5% and Open-Comet gives an accuracy 11.5%. CNN model trained with the original comet assay's data gives lower accuracy if compared to transfer learning model that is trained using other dataset (ImageNet), which proves that kernel learning as a feature extractor in Convolutional Neural Network could be done using other dataset with no correlation to the target dataset [16].

4 Conclusion

In this paper, we have classified comet assay images from buccal mucosal cells with super tiny image dataset using several machine learning-based classification models. The CNN model trained with the original comet assay's data gives lower accuracy than the transfer learning model trained with other dataset. Kernel learning as a feature extractor in CNN could use features obtained from learning other dataset at the target dataset, so transfer learning could be used at cases where the target dataset size is too small. Our transfer learning model has the highest accuracy compares to other methods. Our transfer learning model gives an accuracy 70.5%, while CNN-SVM gives 62.3% and ordinary CNN gives 63.5% and Open-Comet gives an accuracy 11.5%. Our transfer learning model is promising for comet assay for buccal mucosa cell and super tiny dataset.

References

1. Collins, A.R., Oscoz, A.A., Giovannelli, L., Brunborg, G., Gaiva, I., Kruszewski, M., Smith, C.C.: The comet assay: topical issues. Mutagenesis **23**, 143–151 (2008)
2. Kumaravel, T.S., Vilhar, B., Faux, S.P., Jha, A.N.: Comet assay measurements: a perspective. Cell Biol. Toxicol. **25**, 53–64 (2009)
3. Sreelatha, G., Rashmi, P., Sathidevi, P.S.: Automatic detection of comets in silver stained comet assay images for DNA damage analysis. In: 2014 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC) on Proceedings, pp. 533–538, IEEE, Guilin, China (2014)
4. Kiziltan, E., Yurtcu, E.: Semi-automatic scoring tool for comet assay. In: 2015 IEEE 15th International Conference on Bioinformatics and Bioengineering (BIBE) on Proceedings, pp. 1–3. IEEE, Belgrade, Serbia (2015)
5. Mondal, N.K., Bhattacharya, P., Ray, M.R.: Assessment of DNA damage by comet assay and fast halo assay in buccal epithelial cells of Indian women chronically exposed to biomass smoke. Int. J. Hyg. Environ. Health **214**, 311–318 (2011)
6. Matsugu, M., Mori, K., Mitari, Y., Kaneda, Y.: Subject independent facial expression recognition with robust face detection using a convolutional neural network. Neural Netw. **16**(5–6), 555–559 (2003)
7. Dewa, C.K., Afiahayati.: Suitable CNN weight initialization and activation function for Javanese vowels classification. Proc. Comput. Sci. **144**, 124–132 (2018)
8. Kurniawan, J., Syahra, S.G.S., Dewa, C.K., Afiahayati.: Traffic congestion detection: learning from CCTV monitoring images using convolutional neural network. Proc. Comput. Sci. **144**, 291–297 (2018)
9. Adam, R., Dewa, C.K., Afiahayati.: Recognizing Arabic letter utterance using convolutional neural network. In: 2017 18th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD), pp. 181–186. IEEE, Kanazawa, Japan (2017)
10. Dewa, C.K.: Javanese vowels sound classification with convolutional neural network. In: 2016 International Seminar on Intelligent Technology and Its Applications (ISITIA), pp. 123–128. IEEE, Indonesia (2016)
11. Anarossi, E.: Klasifikasi Kerusakan DNA pada Comet Assay Menggunakan convolutional neural network. Theses. Universitas Gadjah Mada, Indonesia (2018)
12. Bakkouri, Ibtissam; Afdel, K.: Breast tumor classification based on deep convolutional neural networks. In: 2017 International Conference on Advanced Technologies for Signal and Image Processing (ATSIP) on Proceedings, pp. 1–6. IEEE, Fez, Morocco (2017)
13. Cortes, C., Vapnik, V.: Support-vector networks. Mach. Learn. **20**(3), 273–297 (1995)
14. Simonyan, K., Zisserman, A.: Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv, pp. 1–14 (2014)

15. Gyori, B.M., Venkatachalam, G., Thiagarajan, P.S., Hsu, D., Clement, M.: Redox biology OpenComet: an automated tool for comet assay image analysis. Redox Biol. **2**, 457–465 (2014)
16. Yosinski, J., Clune, J., Bengio, Y., Lipson, H.: How transferable are features in deep neural networks. In: 27th International Conference on Neural Information Processing Systems—Volume 2 on Proceedings, MIT Press, Montreal, Canada, pp. 3320–3328 (2014)

The New Approach to Detect Scar Healing

Lukas Peter, Antonino Proto, Dan Kamensky and Iveta Bryjova

Abstract The work is focused on the creation of test equipment to verify correct operation cardiotocograph. To simulate the heartbeat of the fetus and the mother was used relay element driven with electricity for AC switching at two selected frequencies to verify the correlation capabilities of commercial cardiotocography. It was done by several commercial device types with specific results.

Keywords Skin colorimeter · Beta prototype · Burn scars · Measurement of skin colour · Hypertrophic scars · Keloid scars

1 Introduction

Nowadays, there are only a few technical devices that make it possible objectively evaluate one or more aspects of the scar. The most common parameters that have been evaluated in recent years using objective technical methods, are the thickness, relief, yield, area and color to which this work is focused.

A clinical evaluation of skin pathology is required to select a diagnostic device (scars) by a doctor and setting a treatment plan. On this basis, you can choose the appropriate diagnostic devices to help your doctor clarify the correctness of the proposed therapy. In this case, it is mainly about type of scar (hypertrophic, keloid, atrophic) or type of skin lesion.

L. Peter (✉) · A. Proto · D. Kamensky · I. Bryjova
Department of Cybernetics and Biomedical Engineering, VSB-Technical University of Ostrava, Ostrava, Czech Republic
e-mail: lukas.peter@vsb.cz

M. Huk et al. (eds.), *Intelligent Information and Database Systems: Recent Developments*, Studies in Computational Intelligence 830,
https://doi.org/10.1007/978-3-030-14132-5_23

2 Problem Definition

The measurement principle is based on reection red and green light, since oxy-hemoglobin has absorption in the spectral part between 520–580 nm, which is an absorption band of green light. This light is much more absorbed than red, which is reacted to the observer, indicating the possibility of calculating the erythema and melanin index, because the difference between the reflected red and green light are getting the necessary values [1].

Deep burning leaves permanent consequences in the form of scars, which are in 30–70 % hypertrophic. And patients considerably individualize. Hypertrophic scars are elevated, painful, red, stiff and itchy, functional and aesthetically unsatisfactory [2]. Do not exceed the limits of the original wound as opposed to scars keloid, which grows into the neighbourhood. The treatment must be long-term and complex (pressure massage of scars, local scouring with anti-scratch, elastic pressure jerseys, laser therapy). All treatment is always with individual effectiveness.

Clinical evaluation of scarring status and development is most common in burn medicine uses the so-called Vancouver Scar Scale (VSS) to score flexibility, pigmentation, height and vascularization [3]. Clinical evaluation of scars is always influenced by choice the evaluator, i.e., always to some extent subjective. Therefore, objective methods, for example, cutometry that measures the viscoelasticity of scars and their development over time. Thus, the effectiveness of the chosen treatment type can be assessed objectively. Unfortunately, cutometry method is expensive and it can be used in all of cases.

3 State of the Art

Clinical research in [1, 4] has found that the incidence of skin cancer in adulthood is narrow association with ultraviolet radiation in childhood, this dependence was unavoidable to investigate and determine the effect of UV radiation on the color changes of children's skin. Color space CIELAB was used to evaluate colorimetric records. Values were recorded for each season. Each subject was measured in six parts of the body. On left and right legs, left and right hands, forehead and armpits that were measured as a control part of the body, which is not exposed to direct sunlight for most of the time. The results showed that the largest color change in the skin occurred in the area of the forearm and the smallest in the area of the armpits and forehead (this is justified by the fact that a total of 30 subjects had hair across the forehead). The results of the clinical trial are based on preventive measures for the problematic areas of the body, especially during the period summer months.

The research [2] was focused on design development of a device that would be able to diagnose decubitus in its early stage (decubitus in this stage are the best response to therapy). This device is composed of LEDs, photodiodes and microcontrollers. Fast and accurate data collection is provided by the software. The principle is based on optical imaging spectroscopy using 5 LEDs of wavelengths (463.7, 532.1, 596, 657, 949.4 nm) and phototransistors. The principle of measurement is based on the detection of changes oxyhemoglobin (HbO2) and deoxyhemoglobin when applied to skin pressure.

The device was constructed from three prototype parts. The first two served to design circuits for LED control. The third one was to construct the final sensor, which was of a circular shape with a diameter 5 cm. Five LEDs were placed around the circuit, in the center of the phototransistor. This whole sensor was lined with black matte paper to block ambient light signals and remove it possible noise of the detected signal. Individual LEDs are covered with white paper around the circuit so that they do not affect each other. This avoided any possible deviations.

The instrument was tested on the upper limb. First, measurements were made for the bony part of the hand, in the elbow area and then on the palm. The individual measurements are always the same the site was measured with pressure application and subsequently without pressure for 15 s, and for individual's light (hypopigmentation) and dark (hyperpigmentation) pigmentation. The results showed that the device is able to distinguish between colors that are very similar to each other, which was tested during the measurement two almost identical shades of the red spectrum, when the device showed fine variations in quantity reflected light.

Possibility to measure skin pigmentation pigmentation by colorimetric or spectrophotometric or image recognition algorithms is real. Subjective methods of evaluation of color changes of hypertrophic scars after burns trauma are abundantly described in specialized articles published in the world's impact logs. However, the method of detecting skin color in an objective manner doesn't exist.

4 Implementation of the New Solution

Several factors influenced the design of the skin colorimeter.

- Simple control
- Low cost device
- Portable device
- Suitable shape and color for medical technology
- Usability in clinical practice

According to these requirements, a prototype was designed to meet these parameters.

DSC1 (Detection of Scar Colour) was selected as a name for the device. The DSC1 prototype is composed of basic electronic and optical components that are collectively capable of detection different color shades. The primary objective of this release was to test whether low cost optical devices can be used components provide the corresponding intensity of incident light and the light reflecting light.

For the proper DSC1 functional page, it was necessary to determine the basic components that they should have to be included in the beta. From a technical point

Fig. 1 The sensor was connected to LCD display and to the Arduino development board to show values of RGB directly on LCD display in real time

of view, they are an electronic system, an optical system and appropriate software. From the user point of view necessary to provide simple controls and a clear view of the processed data. This DSC1 prototype is important in terms of long-term research goals and provides research of technological requirements needed for successful operation of the final version of the device.

In the framework of the implementation of the first generation of the device, several basic electronic ones were used components color sensor and Arduino UNO development board with the ATMega328 microprocessor, the F/F Arduino jumpers, M/M and LCD (Fig. 1).

4.1 Colour Sensor TSC230

It is a programmable color sensor that can detect any number of colors it works on the level of light intensity conversion on the frequency. It consists of 4 LEDs placed in a square arrangement and so-called RGB fields. This field is located exactly in center of the sensor and is composed of 64 silicon photodiodes. In each of these photodiodes are placed 3 filters, first for detecting red (RED) light intensity, second for blue color detection (BLUE), third for Green (GREEN), and the last 16 photodiodes are without a white filter lighting.

Sensor functions are based on the progressive scanning of individual reflected light photodiodes. The output is a rectangular waveform, the frequency of which is given by the generated current, field a photodiode that is proportional to the intensity of the reflected radiation.

Frequency ($f0$) expressed by sensor detection function is described by the following formula.

$$f0 = fd + (Re.Ee) \tag{1}$$

where $f0$ output frequency, fd is the output frequency for dark conditions, i.e., if $Ee = 0$, which is the result of escaping currents, $Ee(mW.cm_2)$ Is the intensity of the incident radiation and ($Re(kHz.mW/cm_2)$ is the response of the sensor for the wavelength of the given light. Because it is $f0$ directly proportional to the frequency, it is possible to convert this frequency into RGB color model. It is possible to clearly define two possible points—absolutely black, which is described by zero with the coordinates [0, 0, 0] and in this case represents the dark conditions of the constant fd, and the absolute white [255, 255, 255], which is the maximum RGB limit (fw), also called white balance. These points form the RGB range [0–255]. Thanks to these parameters, it is possible to express the relationship between frequencies and RGB.

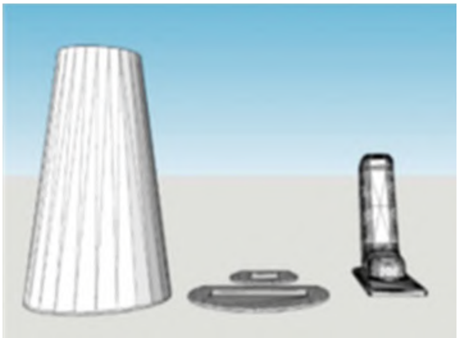


Fig. 2 Whole hardware part was placed into developed plastic case

Fig. 3 Left—view of the sensor cover. Right prototype of DSC1

It is based on the straight line equation. The absolute value of RGB is possible to evaluate from the equation:

$$x = \frac{255(f0 - fd)}{fw - fd} \tag{2}$$

It was necessary to create a sensor shield against the outside illumination (Fig. 2) from a suitable material so as to be suitable for putting the device on the patient's body to make it did not irritate and maintained a constant measurement distance (Fig. 3).

5 Clinical Tests

Verification of the reliability and accuracy of the device has been performed in three clinical cases.

5.1 The First Patient

In the first case, a 28-year-old man was tested who is a hypertrophic scar patient in the dorsal region of the right arm (Fig. 4). 10 measurements were made for the hand scarred for the left hand, thus without disability.

Fig. 4 In the left—dorsal scar on the left arm. In the right—a healthy part of the right arm right

In the case of the first patient, it is clear from the results that the DSC1 showed different results when measuring the healthy and scarred area. The most frequent repeat values for the scar area were R = 179 five times, G = 58 four times, and B = 33 four times as well. The largest deviations for this area were R = 12, G = 9 and B = 3. The smallest deviation for R = 3, G = 1, B = 1. For R, the number did not match, G = 67 three times, B = 38 three times. The greatest deviation for R = 19, G = 9 and B = 29. The smallest deviation for R = 3, G = 1, B = 2.

5.2 *The Second Patient*

In the second case, test was performed on a patient who was at the FNO Burning Center. A man of an older age. It was marked by extensive hypertrophic scarring that was caused by the effects of a high-voltage electrical current, a 15-year-old scar (Fig. 5). Measurement proceeded under constant conditions, temperature 22.8 °C, relative humidity 30.7%. Testing was performed for each affected and healthy part of the body separately and the number of repeats was 10.

The second patient compared 3 scars with surrounding healthy tissues. Measured areas: left shoulder, right thigh and dorsal part of left leg. The measured values for the left-arm hypertrophic scar coincided with R = 156 twice, G = 73 twice and B = 65 twice. The greatest deviations R = 10, G = 18, B = 7, and the smallest measured deviations R = 1, G = 4, B = 1, 3 = 3 times, R = 184 three times, G = 105 three times. The greatest deviations R = 3, G = 3, B = 3. The smallest deviations R = 1, G = 1, B = 1. Right scarred thigh. Match: R = 124 three times, G = 70 five times, B = 67 four times. The greatest deviations: R = 18, G = 7, B = 8. The smallest deviations: R = 1, G = 1, B = 1. The right leg of the thigh. Matches: R = 131 three times, G = 70 five times, B = 63 five times. The greatest deviations: R = 6, G = 3, B = 3. The smallest deviations R = 1, G = 1, B = 1. Dorsal scarring of the left leg. Match: R = 128 three times, G = 72 four times, B = 68 four times. The greatest

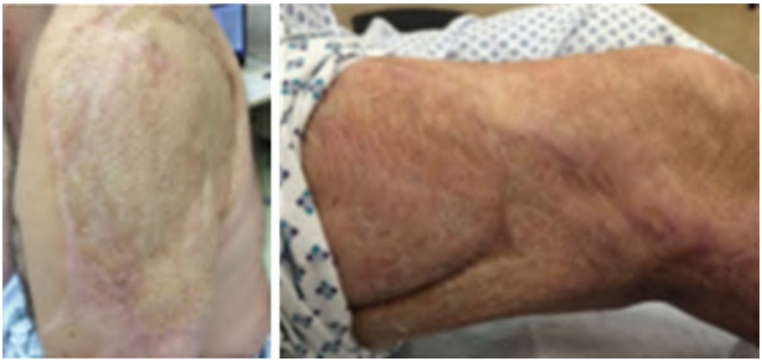


Fig. 5 In the left—scar on the shoulder. In the right a hypertrophic scars on the right thigh

deviations: R = 9, G = 15, B = 33. Smallest deviations: R = 1, G = 1, B = 1. Match: R = 137 three times, G = 76 three times, B = 69 three times. The greatest deviations: R = 20, G = 10, B = 8. The smallest deviations: R = 1, G = 1, B = 1.

5.3 The Third Patient

The last patient did not have a traumatic trauma, involved in this test was due to its higher pigmentation, and thus a good example for comparison with other results. There were clearly identifiable changes in values, especially for parameter R, which is almost doubled. This phenomenon is due to the higher pigmentation of the individual, the more it absorbs the red color, and thus there is a smaller detection of the reflected color. In this case, the most frequently recorded values and their number of repetitions for 10 measurements were as follows, R = 83 four times, G = 53 four times, B = 45 six times. The deviations that were measured at the highest value were R = 4, G = 7, B = 5 and the smallest R = 1, G = 1, B = 2.

6 Discussion

From the measured data, it can be deduced that the device can detect a change when measuring a healthy and scarred area. Testing performed on the non-exposed part showed much smaller variations than scarred parts, which is probably due to the topographic smoother surface of the skin. However, the measurement was influenced by several factors. The greatest variations could be caused by a slight change in the position of the colorimeter and also by the pressure that has been developed at the various times. Last but not least, the results certainly marked the degree of hair. This is shown in particular by the results of the third patient who had the highest number of hair in the area of the left leg.

The analysis of the results also results in the measurement inaccuracies that have already affected the aforementioned factors. To achieve the maximum accuracy of the measurement, it would be necessary to maintain the constant pressure of the colorimeter, which would probably be possible using the spring. In addition, provide shave of the hair of the site and set the exact place of contact with the skin. For further clinical measurements, it would be advisable to follow these conditions in order to determine the exact errors of the device. The total size of the device would be possible in the future by developing a PCB that would disable the Arduino Uno development board and use only its ATMega328 chip, whose area is many times smaller. The last change would be to power the unit directly from the mains (230 V). The frequent replacement of the 9 V battery is unnecessary and unpractical today. An alternative would therefore be to regularly charge the battery from an external power source, that is, from the socket. One of the conditions of use of the device is temperature dependence. The LCD display used with DSC1 is not adjusted for temperatures below 0 °C. If this condition is not met, the crystal may freeze and the LCD display may be compromised. Therefore, it is important to use this prototype above 0 °C.

7 Conclusion

This paper deals with one of the possibilities of objective evaluation of the level of pigmentation hypertrophic scarring after burn injuries. Hypertrophic scars arise especially in terrain at a high degree of burns, where epidermis and dermis are destroyed, and are the most common complication of burn trauma. The prevalence of hypertrophic scarring worldwide is very high.

Evaluating scars using defined rules and scoring scans is inherently subjective. It depends on the experience and skills of the assessing physician, and to some extent limits the validity of this assessment. The result of a physician's subjective view is often limited to determining the level of "scarring" of the scar, and the evaluation results in a hyperpigmented or hypopigmented scar. The appearance of a scar in a subjective assessment can be influenced by ambient light or heat effects, pressure effects of special jerseys, pressure massage or local application of healing agents (creams, ointments, gels) that scars moisturise, soften and help to relieve hypertrophy.

From the acquired theoretical and practical knowledge, a device has been developed that is able to objectively detect the color of the skin and quantify it numerically according to the color space level (R—red, G—green, B—blue). From clinical testing, guidelines (correlation of subjective evaluation with objective skin colorimetry) for a given clinical site will be possible.

Verification of the measuring instrument (skin colorimetry) in clinical practice has clearly demonstrated its functionality and benefit in the objective clinical trial of pigmentation of hypertrophic scars after thermal trauma. Testing of the instrument provided information about the accuracy of the measurement and some deficiencies were evident. The device will also be clinically tested under experimental conditions such as the influence of ambient light and thermal (and local) conditions, drugs applied to the skin, the effect of laser therapy and general therapies. The results will contribute to the development of an extended version of the device, which will be a fully-fledged clinical tool for monitoring the treatment, its success and the setting of optimal therapy with minimal patient burden, all of which are economically acceptable conditions.

It is possible to say that past measurements using the created skin colorimetry have provided clinical opportunity to objective the process of assessing the level of pigmentation of hypertrophic scarring. The possibilities of using this skin colorimeter are not limited to assessing the pigmentation of hypertrophic scarring but can also be used to monitor the effects of pharmacological treatment applied to the skin or practices where skin color identifies the efficacy of the treatment and detects even the slightest advances in therapy.

Acknowledgements The work and the contributions were supported by the project SV4506631/2101 'Biomedicínské inženýrské systémy XII'.

References

1. Zhai, H., et al.: Comparison of tissue viability imaging and colorimetry: skin blanching. Skin Res. Technol. **15**(1), 20–23 (2009)
2. Gaspard, S., Przybycien, T., Siegel, M.: Skin-color-compensated colorimeter for detection and classification of pressure ulcers. In: 2008 IEEE Instrumentation and Measurement Technology Conference Proceedings, IMTC 2008. IEEE, pp. 1542–1547 (2008)
3. Kaartinen, I.: Assessment of Skin Scars in Clinical Practice and Scientific Studies-Introducing Spectrocutometry, a New Objective Method. Tampere University Press, Tampere (2011)
4. Eckhardt, L., et al.: Assessing children's ultraviolet radiation exposure: the potential usefulness of a colorimeter. Am. J. Public Health **86**(12), 1802–1804 (1996)

Correction of the Eye Fundus Blood Vessels Images

Janusz P. Kowalski, Grzegorz Galant, Jakub Peksinski and Grzegorz Mikolajczak

Abstract Analysis of the photographs of the fundus of the eye is a significant diagnostic method used in many medical specialisations. Visible retinal vessels are the only ones in the human body which may be seen with the naked eye. Their condition is the reflection of the health status of various organs and systems. In spite of the appearance of more advanced eye diagnostic methods, the analysis of retinal vessels based on photographs is still being developed. The authors propose increasing of the diagnostic sensitivity through digital filtration of the photographs with the use of a nonlinear order-statistic filter. The obtained results have explicitly indicated a significant increase in readability of the processed images in comparison to the original ones.

Keywords Image processing · Retinal vessels · Fundus of the eye

1 Introduction

The image of the fundus of the eye provides information on the human health condition. Pathological changes visible in the photographs are related to eye diseases or are symptoms of diseases of other organs and systems. Such changes are visi-

J. P. Kowalski (✉) · G. Galant
Department of Computer Science in Medicine, Pomeranian Medical University, Zolnierska 54, Szczecin, Poland
e-mail: janus@pum.edu.pl

G. Galant
e-mail: grzegorz.galant@pum.edu.pl

J. Peksinski · G. Mikolajczak
Faculty of Electrical Engineering, West Pomeranian University of Technology, Sikorskiego 37, Szczecin, Poland
e-mail: jakub.peksinski@zut.edu.pl

G. Mikolajczak
e-mail: grzegorz.mikolajczak@zut.edu.pl

M. Huk et al. (eds.), *Intelligent Information and Database Systems: Recent Developments*, Studies in Computational Intelligence 830,
https://doi.org/10.1007/978-3-030-14132-5_24

ble in the choroid, the vitreous humour and the retina. The source may be a local eye disease or, for instance, a distant inflammation or cancer. The analysis covers the morphology of the blood vessels of the retina and the choroid as well as the optic nerve disc. Due to its structure and function, the retina enables ophthalmic, neurological, laryngological, internal medicine and haematological diagnosis.

Analysis of the retinal vessel morphology, occurrence of retinal artery spasms, retinal arterial embolism, retinal vein thrombosis, retinal vessels inflammation, retina or choroid angiomatosis are only examples of the analysis of the vessels of the fundus of the eye. The causes, on the other hand, may be e.g. atherosclerotic lesions, hypertension, poisonings, inflammations or vegetative disorders [1].

A range of neurological lesions in the brain and the spinal cord first reveal their symptoms in the retina, e.g. Parkinson's disease, Alzheimer's disease [2], ischaemic stroke [3] (microvascular changes suggest a possibility of recurring vascular episodes following an ischaemic stroke). Thus, examinations of the fundus of the eye are also carried out with a view to diagnosing neurological diseases. Many authors point to retinal vessels as the only source of knowledge about changes in small blood vessels of muscles, skin or microvascular structures of the brain. Direct evaluation of small blood vessels of the brain is currently not possible. Nevertheless, it is the evaluation of retinal vessels that provides information on the condition of brain vessels [4].

Pathological changes caused by diabetes—diabetic retinopathy—are clearly visible in the image of the fundus of the eye. This enables determination of the stage of the disease. On the other hand, the appearance of drusen may suggest an early stage of AMD.

The above mentioned examples of using the analysis of the fundus of the eye are only some of the areas where this diagnostic method is applied. They show how significant it is as regards diagnostics in many medical specialisations. An important element of this analysis is evaluation of the health condition based on photographs. The appearance and development of other diagnostic methods, like e.g. optical coherence tomography or optical coherence tomography angiography, have not resulted in abandonment of diagnostics based on photography. Photography is an alternative for more expensive examinations such as coherence tomography. The data concerning renal imaging modality in [5] indicates a quantitative advantage of the use of renal fundus photography in examinations of retinal vascular parameters.

2 Proposed Method

Analysis of images and recognition of the elements within them, particularly performed automatically in decision support systems, are conducted with the assumption that elements of the image with identical colour and brightness belong to the same class. Disturbances in the uniformity of illumination throughout the image, causing regional variations in brightness, may cause difficulties in distinguishing of its elements or even be a source of false interpretation.

The phenomenon of uneven image illumination can be observed in the photographs of the fundus of the eye. The distribution of light intensity depends on the position of the light source in relation to the patient's eye. This distribution is also influenced by the shape of the fundus of the eye, which can be compared to a semi-sphere, and a small opening of the pupil (despite its dilation before the examination) through which the surface under examination is illuminated. Shadows in poorly lit areas make it difficult to analyse the shape and course of blood vessels. In addition, the analysis is hampered by the similar colour of blood vessels and the background on which they occur. Adjustment of the light, especially with colour correction of the photography elements, is a way to increase the readability of the photographs of the fundus of the eye.

A method that enables a significant reduction of global illumination unevenness is the so called background illumination modelling. The background is a smooth image with the same unevenness of illumination as the processed image. It is assumed that objects in the image are represented in the domain of higher frequencies, while low frequency variable illumination is manifested mainly in low frequencies [6]. Depending on the characteristics of the recording device, the background image is subtracted from the source image or the source image is divided by the background.

The study used the rank-based levelling method. This is due to the properties of a photograph: the variation in background illumination is so large that it cannot be modelled with a relatively low degree polynomial, whereas the background level is, depending on the region, brighter or equal to the level of the objects to be separated (blood vessels).

Background separation was performed with the use of a nonlinear order-statistic filter. As the aim was to reflect the illumination, a filter was used the response of which is 100th percentile of a ranked set of the values of the pixels contained in the area encompassed by the filter's spatial mask. It can be expressed by:

$$g(x, y) = \mathrm{MAX}[v\colon v = \{v(r)\}, r = 1, 2, \ldots, m \times m] \tag{1}$$

where $g(x, y)$ is the resulting value of the central pixel with the coordinates (x, y) of the area encompassed by the filter mask of size $m \times m$, and $\mathbf{v}$ is the one-dimensional vector of the pixels illumination, sorted in ascending order according to their values.

Estimation of the illumination level for a pixel is a function of illumination and reflectivity. Both quantities are variables of a local nature. Since this chosen method takes into account a local variation, order-statistic filtering is locally adaptive. This has resulted in more efficient modelling of the local light level. The filtration process was carried out for two arbitrarily selected mask sizes: 3×3 and 5×5.

The result of the filtration is the distribution of background illumination g(x, y). In order to correct the disturbed uniformity of illumination throughout the image, the distribution obtained was subtracted from the source image. Calculations were done on colour images decompressed into RGB space. Separate, independent filtration in each RGB channel enabled obtaining of the distribution of background illumination for each component, which enabled independent corrections for each of the components. The last step of the processing was gamma correction.

3 Experimental Results and Discussion

The results of the performed operations are presented in Figs. 1, 2, 3, 4, 5 and 6.

Figures 1–3 show the correct image of the fundus of the eye. The subsequent ones are an example of processing the photographs where pathological changes are visible. The lesions occurring in the recorded images were not masked with the proposed processing method, but emphasized.

Figure 4 presents druzen, being senile and sclerotic lesions, which do not impair the visual function but may cause damage to the retina itself. They are visible as small shiny spots under the retinal vessels. In processed photographs, their image is clearer.

The blood extravasation is visible in the photographs as a dark area, masking the blood vessels within its range (Fig. 5). After processing, the image of blood vessels hidden in the original photograph by the extravasation was shown clearly and the

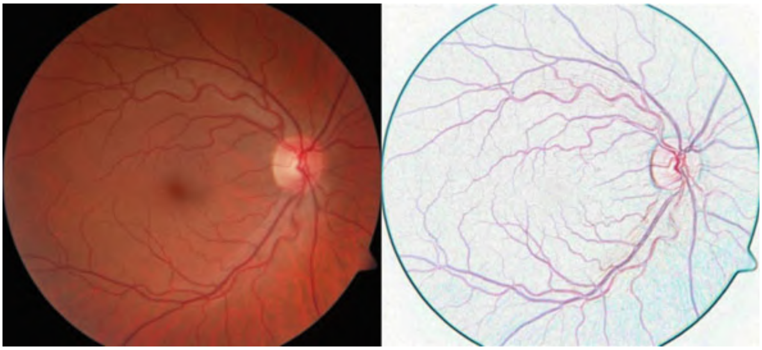


Fig. 1 Images of the fundus of the eye: original (*on the left*), processed (*on the right*)

Fig. 2 Images of the fundus of the eye: original (*on the left*), processed (*on the right*)

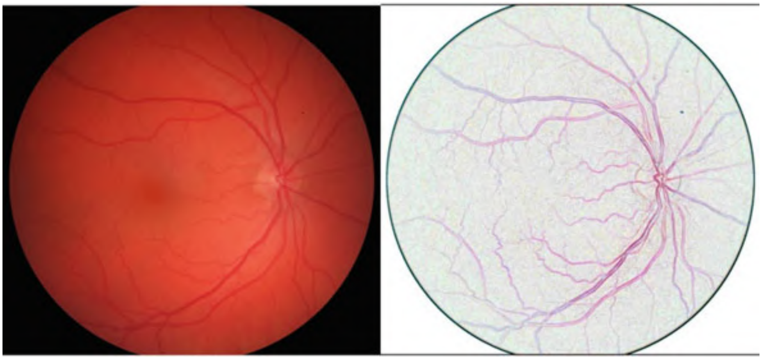


Fig. 3 Images of the fundus of the eye: original (*on the left*), processed (*on the right*)

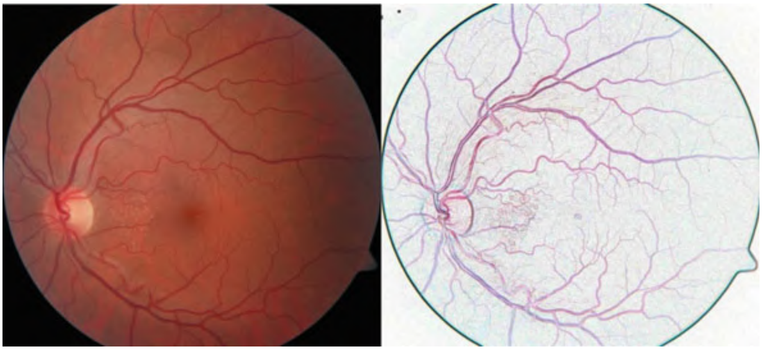


Fig. 4 Images of the fundus of the eye: original (*on the left*), processed (*on the right*)

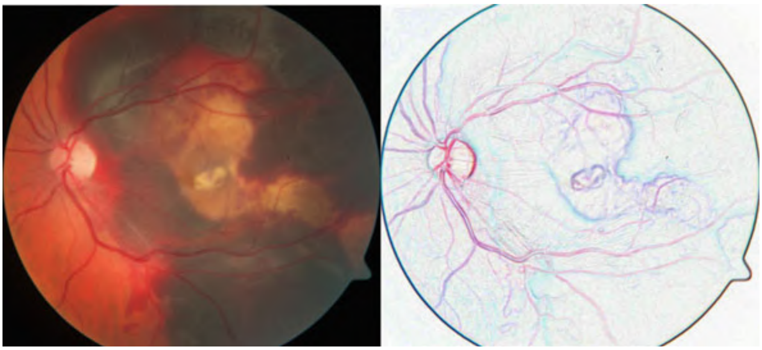


Fig. 5 Images of the fundus of the eye: original (*on the left*), processed (*on the right*)

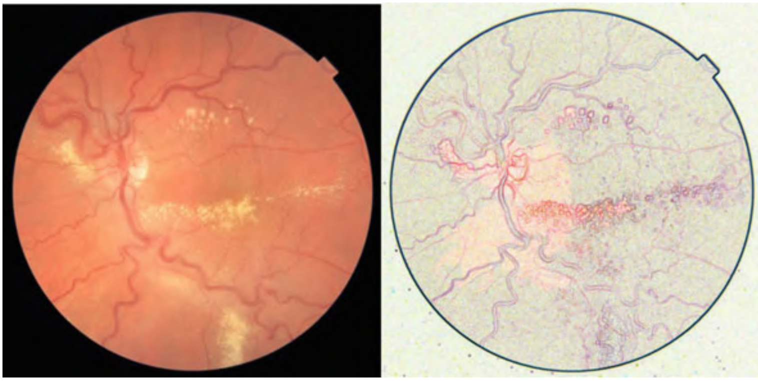


Fig. 6 Images of the fundus of the eye: original (*on the left*), processed (*on the right*)

extravasation border was marked. Another example is lipid degeneration of the retina, which was also emphasized in the processed image (Fig. 6).

As the objective was to improve the diagnostic capability, verification involved assessment of the difference in readability of source and processed images performed by physicians. In each case, regardless of the source of the original image—current or archival, they confirmed a significant increase in the readability of the results obtained.

4 Conclusions

The use of colour fundus photography for long-term follow-up of retinal nerve fibre layer cleavages is described by Hsia et al. [7]. Photography is used except for basic ophthalmic examinations, optical coherence tomography and visual field analysis to investigate the structural and functional characteristics and change of the retinal nerve fibre layer cleavages in glaucoma patients and suspects. Marupally et al. [8] use digital colour fundus photographs in diagnostics of diabetic retinopathy, which is one of the main causes of blindness.

The aim of the study of Benson et al. [9] was to describe dynamic colour change in retinal vessels due to externally applied ocular pressure in a 6-week-old infant with lipemia retinalis secondary. The research was conducted with the use of colour photographs of the fundus.

The accurate segmentation of the vessel is the prerequisite step for automated or computer-aided diagnosis systems. A clear image of all central and peripheral vessels is essential for proper functioning of the HDSS system. Jiang et al. [10] describes the problem of segmentation of the retinal vessel tree from full-size images. The same problem, also in the context of machine learning, resulting from the ambiguity between the image of thick and thin vessels, is indicated by Yan et al. [11] or Yang et al. [12].

Another example of using digital photos of the eye is examination of the retinal vessel calibre at delivery and one-year post-partum of women who have had pre-eclampsia during pregnancy. The observed changes may reflect a permanent, long-term microvascular dysfunction [13].

Cheung et al. [14] point to limitations of current neuroimaging technology direct in vivo visualization of the cerebral small vessels. Nevertheless, thanks to similar embryological origin, anatomical features, and physiologic properties of the cerebral small vessels and the retinal vessels, the latter reflect condition of the cerebral vessels. The authors use this dependence by applying ocular fundus photography in neurological diagnostics. The research conducted by the authors suggests that the image of retinal vascular changes seen on fundus photography is strictly related to stroke and dementia. McGrory et al. [15] describes inconsistent pathological changes observed in fundus camera imaging as regards vessel caliber, tortuosity, and fractal dimension, which are typical for dementia. Perez et al. [16] not only describes a huge potential of retinal photography for neurology or emergency medicine, but also points to the possibilities of using a digital form of photographs in telemedicine.

The significance of the image of the fundus of the eye in the work of emergency department physicians is described by Bruce et al. [17, 18]. Photographs are recorded without the use of a dilation agent, which results in worse illumination of the fundus. Photographs are placed on the electronic medical record for emergency physician review. The image obtained is the basis for diagnosis. Simultaneously, due to the nature of the work of emergency physicians, nonmydriatic ocular fundus photography is a very frequently used method as an alternative to mydriatic ocular fundus photography.

Current world literature shows that, despite the development of new, very advanced diagnostic methods, such as optical coherence tomography, the analysis of blood vessels based on photographic images of the fundus of the eye, as a diagnostic method, is used and develops in many medical specialisations. The proposed method of image processing increases the sensitivity of blood vessel analysis based on a photograph, regardless of its use by specialists in various areas of medicine.

The evaluation of the proposed method includes a medical and economic aspect. Resulting from the emphasizing of the peripheral parts of the retina, it is also possible to assess this part of the fundus of the eye, which is especially important in the diagnosis of diabetic retinopathy and other diseases causing retinal ischemia.

Taking into account the criterion of inconvenience for a patient, it should be emphasized that the analysis of the fundus of the eye based on photographic images is a non-invasive diagnostic method. A patient is given only a medication dilating the pupils, after which a photograph is taken without any additional interference in the body. In some cases, the photo is taken even without the use of a dilating medication [17].

In some examinations of the fundus of the eye, contrast is additionally used as one of the ways of clearer imaging of blood vessels, but then such tests become invasive. What cannot be underestimated is the fact that bad tolerance for contrasting agents has become more common. The administration of a contrasting agent also involves additional costs of a pharmaceutical. The proposed method does not require the administration of a contrasting agent and, thus, it is neither burdensome for the body nor involves any additional costs, at the same time being a more sensitive diagnostic tool.

Another medical aspect of the proposed method is the ability to assess the progress of the disease on the basis of a comparative analysis of current and archival images (patients oftentimes bring their own documentations of the fundus of the eye created years ago). The source of the image (type of diagnostic apparatus) and the time of its creation do not matter; algorithm is not a part of the diagnostic system, it operates as an independent software.

Previous studies have demonstrated the versatility of the proposed method in relation to different types of eye fundus images. Further studies are focused on the assessment of the impact of the proposed method on the increase of diagnostic sensitivity in particular types of diseases.

References

1. Bryk, E.: Dno oka w chorobach pozaukładowych. PZWL, Warszawa (1989)
2. London, A., Benhar, I., Schwartz, M.: The retina as a window to the brain—from eye research to CNS disorders. Nat. Rev. Neurol. **9**, 44–53 (2013)
3. De Silva, D.A., Manzano, J.J.F., Liu, E.Y., Woon, F.-P., Wong, W.-X., Chang, H.-M., Chen, C., Lindley, R.I., Wang, J.J., Mitchell, P., Wong, T.-Y., Wong, M.-C.: Retinal microvascular changes and subsequent vascular events after ischemic stroke. Neurology **77**(9) (2011)
4. Abdelhak, A., Hübers, A., Böhm, K., Ludolph, A.C., Kassubek, J., Pinkhardt, E.H.: In vivo assessment of retinal vessel pathology in amyotrophic lateral sclerosis. J. Neurol. **265**(4), 949–953 (2018)
5. Dumitrascu O.M., Qureshi T.A.: Retinal vascular imaging in vascular cognitive impairment: current and future perspectives. J. Exp. Neurosci. **12** (2018)
6. Parker, J.R.: Algorithms for image processing and computer vision. Wiley Computer Publishing, New York (1997)
7. Hsia, Y., Su, C.C., Wang, T.H., Yang, C.M., Huang, J.Y.: Long-term follow-up of retinal nerve fiber layer cleavages in glaucoma patients and suspects. Graefes Arch. Clin. Exp. Ophthalmol. **256**(10), 1945–1952 (2018)
8. Marupally, A.G., Vupparaboina, K.K., Peguda, H.K., Richhariya, A., Jana, S., Chhablani, J.: Semi-automated quantification of hard exudates in colour fundus photographs diagnosed with diabetic retinopathy. BMC Ophthalmol. **17**(1), 172 (2017)
9. Benson, M., Wong, R.W., Young, R.C., Gibson, J.B., Fan, Y., Harper, C.A.: Dynamic retina vessel color change during ocular compression in severe lipemia retinalis. Retin. Cases Brief. Rep. **12**(Suppl 1), S98–S101 (2018)
10. Jiang, Z., Zhang, H., Wang, Y., Ko, S.B.: Retinal blood vessel segmentation using fully convolutional network with transfer learning. Comput. Med. Imag. Graph **68**, 1–15 (2018)
11. Yan, Z., Yang, X., Cheng, K.T.: Joint segment-level and pixel-wise losses for deep learning based retinal vessel segmentation. IEEE Trans. Biomed. Eng. **65**(9), 1912–1923 (2018)
12. Yang, Y., Shao, F., Fu, Z., Fu, R.: Blood vessel segmentation of fundus images via cross-modality dictionary learning. Appl. Opt. **57**(25), 7287–7295 (2018)
13. Soma-Pillay, P., Pillay, R., Wong, T.Y., Makin, J.D., Pattinson, R.C.: The effect of pre-eclampsia on retinal microvascular caliber at delivery and post-partum. Obstet. Med. **11**(3), 116–120 (2018)

14. Cheung, C.Y., Chen, C., Wong, T.Y.: Ocular fundus photography as a tool to study stroke and dementia. Semin. Neurol. **35**(5), 481–490 (2015)
15. McGrory, S., Cameron, J.R., Pellegrini, E., Warren, C., Doubal, F.N., Deary, I.J., Dhillon, B., Wardlaw, J.M., Trucco, E., MacGillivray, T.J.: The application of retinal fundus camera imaging in dementia: a systematic review. Alzheimers Dement (Amst) **6**, 91–107 (2016)
16. Pérez, M.A., Bruce, B.B., Newman, N.J., Biousse, V.: The use of retinal photography in nonophthalmic settings and its potential for neurology. Neurologist **18**(6), 350–355 (2012)
17. Bruce, B.B., Thulasi, P., Fraser, C.L., Keadey, M.T., Ward, A., Heilpern, K.L., Wright, D.W., Newman, N.J., Biousse, V.: Diagnostic accuracy and use of nonmydriatic ocular fundus photography by emergency physicians: phase II of the FOTO-ED study. Ann. Emerg. Med. **62**(1), 28–33 (2013)
18. Bruce, B.B., Biousse, V., Newman, N.J.: Nonmydriatic ocular fundus photography in neurologic emergencies. JAMA Neurol. **72**(4), 455–459 (2015)

Part V
Applications of Intelligent Information Systems

Implementation of Decision Tree Algorithm on Game Agent of First Aid Educational Game

Gunawan, Asep Nurhuda and Reza Andrea

Abstract Research on first aid educational game is a research development of Drag and Drop Game type games designed to provide learning how to deal with injured patients. In this game players must provide treatment to patients affected by various injuries. The shuffle-shuffling algorithm was applied in this study, with the aim of making the composition of injured patients invariably scrambled and the game not monotonous and boring. AI technology (artificial intelligence) is also applied to this research. Using the Decision Tree modeling method, games that are built will have game agent characters that will accompany the child to play like a patient, so the game feels more had a real interaction.

Keywords Educational game · First aid · Game agent · Decision tree

1 Introduction

Education game is one of the implementations of computer science whose development has been very rapid. An educational game is also an educative application, meaning it can be used as a learning medium where the process is done with the concept of learning while playing.

The learning process using a book or conventional learning stage sometimes makes students less interested, because the book only displays instructional media in the form of text, and learning media in the book cannot be visualized so it looks

Gunawan
STIMIK, Balikpapan 76144, Indonesia
e-mail: gunawan@stmikbpn.ac.id

A. Nurhuda
STMIK Widya Cipta Dharma, Samarinda 75123, Indonesia
e-mail: acep.noor@wicida.ac.id

R. Andrea (✉)
Politeknik Pertanian Negeri Samarinda, Samarinda 75117, Indonesia
e-mail: reza@politanisamarinda.ac.id

M. Huk et al. (eds.), *Intelligent Information and Database Systems: Recent Developments*, Studies in Computational Intelligence 830,
https://doi.org/10.1007/978-3-030-14132-5_25

boring (not moving, black and white), this can cause children to feel bored and their interest in learning is reduced [1].

The education of first aid has been ignored in this era of technology. The modern knowledge of first aid tends to be absent because children's time is consumed by playing smartphone games from international developers. The presence of the education game is a solution that can bring knowledge and learning of first aid into a game. This encourages children to keep learning but in the form of playing games.

Implementing educational games as a learning medium can create a fun new atmosphere in the learning process. Picture and sound make learning process become exciting. Therefore, this study is conducted to build first aid education game, a game that teaches children how to provide first aid to injured patients. The game is built in the type of drag and drop game designed to sharpen memories. Players can use the first aid tools available in the game scene to provide first aid treatment to the accident victims. In learning, it requires an instructor or teacher as a student companion to learn. Thus, the role of companion will be replaced by the game agent [2, 3]. Game agents are intelligently controlled game characters using decision tree logic. A decision tree can be used in designing AI (artificial intelligence) for games.

The decision tree is known as a technique for modeling phenomena or event-based conditions. The decision tree is what will make the game characters that accompany children while playing can think for themselves according to the condition of the game. Characters can give a happy, angry, or sad response according to the child's play style. This research expects the first aid game to be a fun first aid educational game and it will not make the children bored when playing. It can also be a playing tool for children that makes them able to learn anywhere through the media game on android smartphone.

1.1 Problems Identification

The problems identified in this research are:

1. There is no previous research on how to build drag and drop type education game on first aid.
2. There is not yet first aid educational game that has a companion for the child while playing in the form of a game agent with AI (Artificial Intelligence) [4–6].

1.2 Purpose of the Research

The purpose of this research is:

1. Create a drag and drop game about first aid.
2. Implement the Decision Tree logic in the form of a frame model to be applied in the game character agent.

2 Related Works

Research on games with the same technique has been done before, namely:

1. Game Tree Search Based on Nondeterministic Action Scripts in Real-Time Strategy Games [7]
2. The Tracking Technique of the Object Position in a Game Called "Find Me!—Bumi Etam" [8]
3. The Tracking Technique of Arena in a Game Called "Find Me!—Earth Etam" with Random Shuffle Technique [9]

In Barriga, et al. research, Support vector machines (SVMs) are an emerging and powerful technique in coping with classification problems. However, a lack of rule generation is a weakness of the SVM model, especially in analyzing sporting results. This investigation developed a hybrid model integrating the SVM technique and a decision tree approach (HSVMDT) to predict the results of basketball games, and to provide rules to aid coaches in developing strategies. The HSVMDT model employed the unique strength of SVM and decision tree in generating rules and predicting the outcomes of games. With predicted outcomes of games, and rules yielded from the HSVMDT model, coaches can easily and quickly learn essential factors increasing the chances to win games.

In two game research Find Me!, the player must find where the funny character object "Nguk-Nguk" is hiding. The hiding arena of this cute character cannot be arranged in a regular level sequence, as this will definitely make the game monotonous and boring. Arranging the order of the arena to be faced by players in each level, should be done by random shuffle randomization technique, with the aim of making the game more challenging.

Differences in this study, the game to be built using the decision tree modeling method, games that are built will have game agent characters that will accompany the child to play like a patient, so the game feels more had a real interaction.

2.1 Research Stages

This research is conducted through the stages of multimedia development as follows:

1. Creating the concept of gameplay and the way the algorithm is applied, as well as collecting game-making materials.
2. The development process of educational game with multimedia system development method, from designing to the assembly stage of decision tree implementation into game system.
3. Installing educational game apps into children's smartphones and elementary school teachers.
4. Testing the game to produce satisfactory percentages.

2.2 Decision Tree of Game Agent

The decision tree is a classification method that uses a representation of a tree structure in which each node represents an attribute. Its branch represents the value of the attribute, and the leaf represents the class [10, 11]. The topmost node of the decision tree is called root. Decision tree is the most popular classification method used. In addition to its relatively rapid development, the results of the built model are easy to understand. In the decision tree there are three types of nodes [12], namely:

1. Root Node is the topmost node.
2. An Internal node is the branch node.
3. Leaf node or terminal node is the final node.

From the tree in Fig. 1, the root node is "should they be given plaster". This is the first condition of handling injuries in this game. If this condition is True (plaster handling = True), then the path will lead to leaf node "smile" which means game agent will give smile expression. Whereas, if the condition is False, then the path will go to internal node "should they be compressed". This rule affects the flow of the expression of the patient (game agent). In leaf node, the game agent will give a happy expression which means all the wounds have been healed.

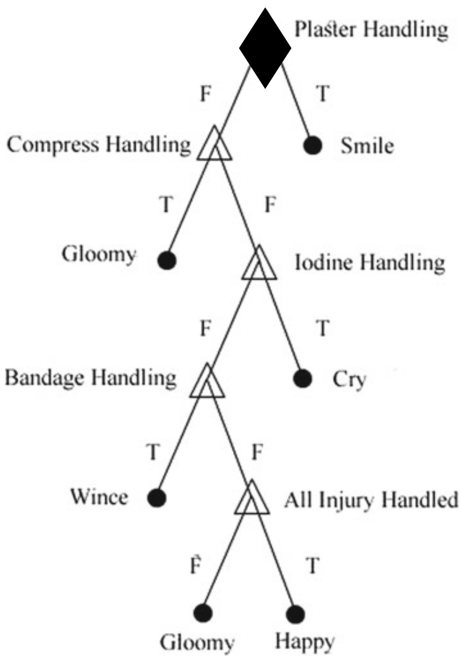


Fig. 1 Logic model of decision tree in the patient: black diamond symbol is root node, triangle symbol is internal node and black bullet is leaf node

3 Results and Analysis

"First Aid Care" is a Drag and Drop educational game which is included in the category of game puzzles that can make players to play while learning in a relaxed manner. This game is easy to learn or play by the players. The game has simple rules, and it displays content designed for its users for all ages, especially children. This game has four menu options in the scene opening that is the introduction of first aid kit, how to play, play, and the exit menu. Concept is assembled on the interface of Fig. 2. This Educational game has rules in which players must treat injured patients by dragging and dropping the first aid kit that has been played by the player. This Educational game has rules in which players must treat injured patients by dragging and dropping the first aid kit that has been played by the player. This concept is assembled on the interface of Fig. 2.

In the first aid kit introduction menu, the players will learn about the function of first aid kit with its name and pictures (Shown in Fig. 3).

Play scene is a scene where the user plays by giving first aid to the injured patient in accordance with the instructions given by the doctor. The player must complete five levels. If the user has finished treating the injured patient, then the user can proceed to the next level. If the user has completed all five levels, they will enter the award scene and players can repeat the game from the first level, with each level being randomized to the number and different types of treatment—so that the game is not monotonous (Shown in Fig. 4).

Fig. 2 Opening scene first aid

Fig. 3 First aid learning scene

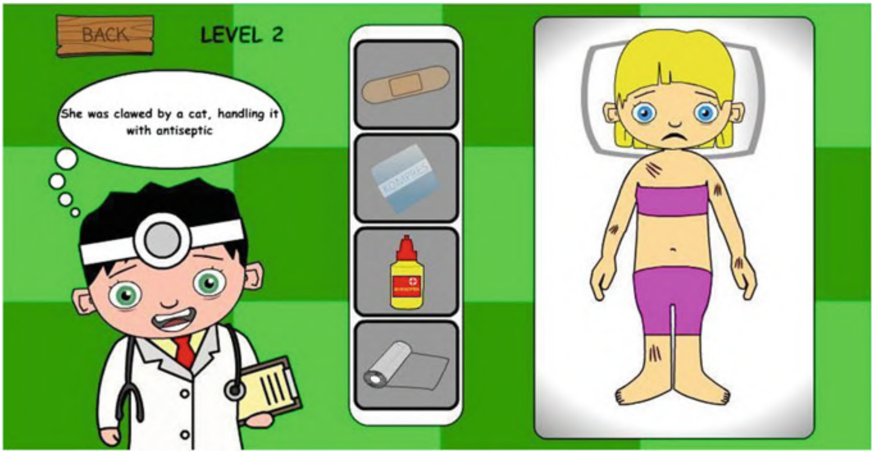


Fig. 4 Gameplay first aid educational game

3.1 Assembly Rule of Decision Tree

The Decision Tree logic model (Fig. 1 on the previous page) is applied as a pattern of expression of the patient character on the game. There will be expression given to the player who has treated the injured patient (Fig. 5).

Decision tree design expression of patient, it implemented to pseudo code below

```
Function Patient_DTree(){
          If (plaster_handling == true){
                 Patient_smiley = "smile";
          }Else if (compress_handling == true){
```

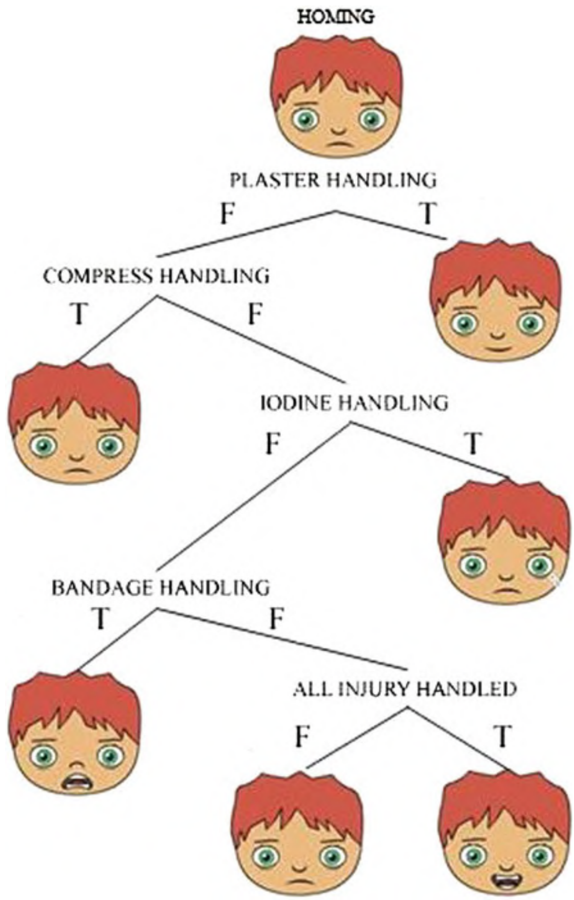


Fig. 5 Expression of the treated patient

```
            Patient_smiley = "gloomy";
        }Else if (iodonie_handling == true){
            Patient_smiley = "cry";
        }Else if (bandage_handling == true){
            Patient_smiley = "wince";
        }Else if (all_injurry_handled == true){
            Patient_smiley = "happy";
            GotoSceneandPlay("Complete_scene",1);
        }Else{
          Patient_smiley = "gloomy";
        }
}
```

It can be seen in Fig. 6, that the patient can give a sad, happy, wincing, crying, and happy expression at the time the injury is treated. Pseudo code of the decision tree logic will make this expression appear based on the action of the player. Figure 6 shown, at the last internal node, if all injury handle is true patient will happy and the

Fig. 6 The scene interface of completing the treatment

game will show completed treatement scene (see Fig. 6), else patient keep gloomy and player cannot go to the next level.

All of these appearances and expressions are made with the intention of making the game interface more interacting with players who are still children, so the players of this educational game feel accompanied by the presence of game agents.

3.2 Beta Testing

Beta testing is a live app test in an environment that cannot be controlled [13]. The testing was conducted using a simple questionnaire filled out by teachers and elementary school students. In this study, the beta testing was conducted on 2 teachers and 8 students who had to answer two questions referring to the gameplay and the game agent role. Questionnaires are made as simple as possible so that children can also fill it out.

From the result of beta test in Table 1, it can be calculated for answer "less" with 0, for answer "satisfactory" with 1, and answer "good" with 2. After that, the mean percentage of respondent value can be calculated (1).

$$\bar{X} = \frac{17 \times 2 + 3 \times 1 + 0 \times 0}{20 \times 2} \times 100 = 92.5\% \quad (1)$$

Based on the percentage calculation, the percentage is 92.5%. It means the first aid educational game is acceptable as seen from the percentage obtained, above the minimum percentage of 50% (good), and close to 100% (very good).

Table 1 Result of beta testing

Question	Respondent's questions			Total respondents
	Good	Satisfactory	Poor	
How is this educational game?	8	2	0	10
How is the characters of patient and doctor in the game?	9	1	0	10
Total answer	17	3	0	**20**

4 Conclusion

The results of this research, first aid educational game has been built. It is an educational game that teaches how to treat an injured person in the form of drag and drop game. The application development is done according to the stage of multimedia development, starting from concepting the gameplay to the beta testing.

Present the game agent as a companion character when children are playing. The decision tree applied to the game agent makes a cute expression and it can interact based on the action-reaction logic of the player's style. This is what makes the educational game more interactive to children. We suggest for future work, first aid educational game can use another method is like behavior tree, finite state machine, or combination decission tree and finite state machine to make the game agent more interactive.

References

1. Wijayanti, S., Andrea, R.: *K*-means cluster analysis for students graduation: case study: STMIK Widya Cipta Dharma. In: Proceedings of the 2017 International Conference on E-commerce, E-Business and E-Government (pp. 20–23). ACM 2017
2. Andrea, R., Kopel, M.: Design and development of "battle drone" computer-based trading card game (CTCG). In: International Conference on Multimedia and Net- work Information System (pp. 574–585). Springer, Cham (2018)
3. Andrea, R., Akbar, R.I., Fitroni, M.: Developing battle of etam earth game agent with finite state machine (FSM) and sugeno fuzzy. ICCS Proceeding **1**(1), 184–187 (2014)
4. Charlier, N., De Fraine, B.: Game-based learning as a vehicle to teach first aid content: a randomized experiment. J. Sch. Health **83**(7), 493–499 (2013)
5. Charlier, N.: Game-based assessment of first aid and resuscitation skills. Resuscita- tion **82**(4), 442–446 (2011)
6. de Urturi, Z.S., Zorrilla, A.M., Zapirain, B.G.: Serious game based on first aid education for individuals with autism spectrum disorder (ASD) using android mobile devices. In: IEEE Intelligent Systems (pp. 223–227). IEEE (2011, July)
7. Barriga, N.A., Stanescu, M., Buro, M.: Game tree search based on nondeterministic action scripts in real-time strategy games. IEEE Trans. Games **10**(1), 69–77 (2018)

8. Andrea, R.: Teknik Pengacakan Posisi Objek “Find Me!-The Game”. “object position randomization technique “Find Me!—The Game”. In: Proceeding SENAIK, pp. 301–303. 1 Nov 2013, Samarinda, Indonesia (2013)
9. Andrea, R.: Game arena randomization technique “Find Me!—Bumi Etam” with the technique of shuffle random. In: Proceeding SENATKOM, pp. 823–827. 23 Sept 2015, Padang, Indonesia (2015)
10. Poole, D.L., Mackworth, A.K.: Artificial intelligence foundations of computational agents. Cambridge University Press, New York (2010)
11. Russell, S.J., Norvig, P.: Artificial intelligence a modern approach, 3rd edn. Pearson Education, Inc., United States of America (2011)
12. Kyaw, A.S. Peters, C., Swe, T.N.: Unity 4.x Game AI programming. Packt Publishing Ltd. Birmingham (2013)
13. Hurd, D., dan Jenuings, E.: Standardized Educational Games Ratings: Suggested Criteria. Longman. London (2009)

A Generic Framework for Cross Domain Recommendation

Muhammad Murad Khan and Roliana Ibrahim

Abstract Cross-domain recommender systems research focuses on recommendation improvement in the target domain based on knowledge gathered from the source domain. Recently, a number of review studies were published, however, they did not provide their reader with a step-by-step guide to conduct cross-domain recommender systems research. In this paper, we present a generic cross domain recommender systems framework which identified different phases and components required to execute a cross-domain recommendation research. The proposed framework is based on analysis of primary studies collected from recently published high impact review studies. It was found that all of the collected primary studies included identified phases, however, each component had multiple options (approaches) which can be applied based on scenarios. The proposed framework can be used for analysis and comparison of existing primary studies, and also explains why cross domain recommender systems approaches are evaluated the same as conventional recommender system approaches. Finally, conclusion and future direction are presented.

Keywords Cross domain · Recommender systems · Knowledge transfer · Framework

1 Introduction

Recommender systems are special software designed to solve information overload problem [12]. Information overload occurs when a user face difficulty to identify required item/object/information etc., because of the complexity of the system. Hence recommender systems attempt to tailor user experience by providing a list of items/objects/information based on user's past experience or search query.

M. M. Khan (✉) · R. Ibrahim
Universiti Teknologi Malaysia, 81310 Skudai, Malaysia
e-mail: muradtariq.tk@gmail.com

R. Ibrahim
e-mail: roliana@utm.my

M. Huk et al. (eds.), *Intelligent Information and Database Systems: Recent Developments*, Studies in Computational Intelligence 830,
https://doi.org/10.1007/978-3-030-14132-5_26

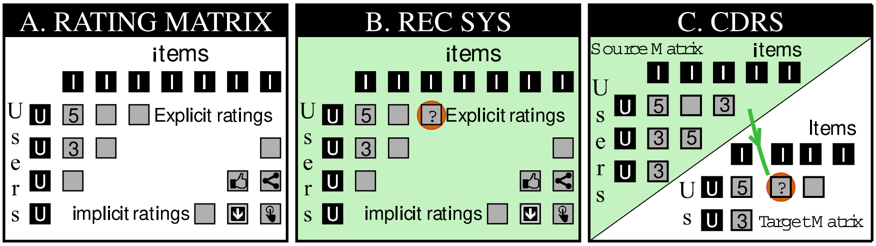


Fig. 1 Rating matrix, rating prediction and cross domain recommendation

Figure 1a represents a rating matrix, which is considered as a building block of a recommender system. It consists of users, items and interactions which can be explicit (numeric: 0–5) or implicit (like, download, etc.). Recommender system (REC SYS) research focuses on predicting a value for an item having no ratings, illustrated in Fig. 1b. whereas, cross-domain recommender systems (CDRS) research focuses on predicting rating in target domain based on knowledge available in the source domain [14], illustrated in Fig. 1c.

This research focuses on different components researchers use to execute cross-domain recommendation and finally propose a cross-domain recommendation framework which can be used for understanding and comparing different cross domain recommender systems approaches. The proposed framework was developed using primary studies of high impact systematic literature reviews conducted by [13, 20].

This paper is divided into five sections. After presenting the introduction, Sect. 2 focuses on existing literature. Section 3 presents the methodology of this research paper and identify the need of CDRS framework. Section 4 presents the cross-domain recommendation framework and finally, Sect. 5 presents the conclusion and future work.

2 Literature Review

New researchers are facing confusion when it comes to cross-domain recommendation [13], hence, it was found appropriate to collect cross domain recommender systems primary studies, identify common phases and components enabling cross-domain recommendation, and based on collected primary studies, propose a generic framework that can simplify understanding of CDRS research.

This study focused on review studies related to CDRS for the collection of primary studies. Existing literature was explored and two review studies were identified which conducted a thorough analysis of CDRS. Both studies [13, 20] were published within the last three years and based their analysis on 94,34 primary studies mentioned in studies [13, 20].

First study [20], analyzed and classified the existing cross-domain recommendation techniques. First, they identified concepts involved in the cross-domain recommendation and later used identified concepts to explain different transfer-learning techniques. They divided transfer learning techniques into two main categories i.e., knowledge aggregation and knowledge linkage. Each category further had three sub-categories. They also proposed multiple future directions such as analysis of context as a domain; utilization of versatile evaluation metrics such as diversity, novelty, and serendipity for illustration of CDRS benefits; development of elicitation tool for target recommendation; and stressed need of real-world datasets.

Second study [13], on the other hand, classified existing research with respect to domain vs user-item overlap; Recommendation tasks vs. user-item overlap; Cross-domain recommendation approaches; evaluation metrics and utilized datasets; and trending recommendation problems. They also proposed multiple future directions such as: enhancing domain similarity by integration of heterogeneous data available on social networks; development of scalable and distributed CDRS algorithms; exploration of relatively new research problems in recommender systems such as risk, adaptively and robustness; and creation of new datasets based on real-world data.

This study relies on primary studies identified in the review studies and analysis methodology is discussed in the next section.

3 Methodology

This study utilizes 94 and 34 primary studies analyzed in [13, 20] respectively. Based on common concepts utilized by each primary study, three phases are identified which are 'domain selection criteria', 'knowledge transfer' and 'recommendation generation'. These phases were present in each cross domain recommender systems primary study, however, execution of each phase varies based on approach/scope of the study.

Variation between studies is grouped with respect to different components that connect each phase for smooth execution of cross-domain recommendation. Different components have multiple implementation options (approaches) which are discussed in the next section, however, for each option, example studies are cited. Example studies come from the collected bank of primary studies.

Based on the used methodology, this study attempts to present a generic cross domain recommendation framework which is explained in detail in the next section.

4 Cross Domain Recommendation Framework

Cross-domain recommendation framework consists of three phases. First phase, "domain selection criteria", consists of three components, that is, "target domain selection criteria", "problem identification" and "source domain selection criteria". Second phase "knowledge transfer" consists of four components. That is "similarity

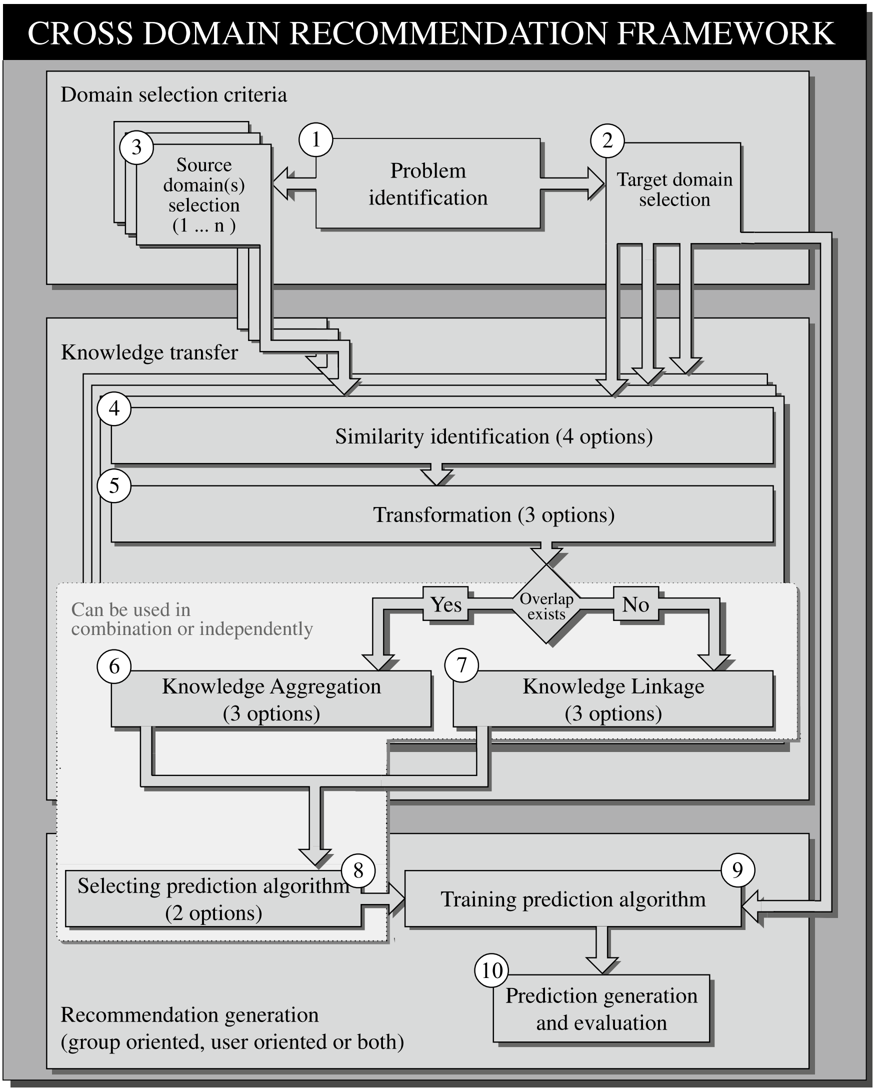


Fig. 2 Cross domain recommendation framework

identification", "interaction transformation", "knowledge aggregation" and "knowledge linkage". Third phase recommendation generation consists of three components. That is, "prediction algorithm selection", "training prediction algorithm" and "recommendation generation". Flow between each phase and different components is illustrated in Fig. 2.

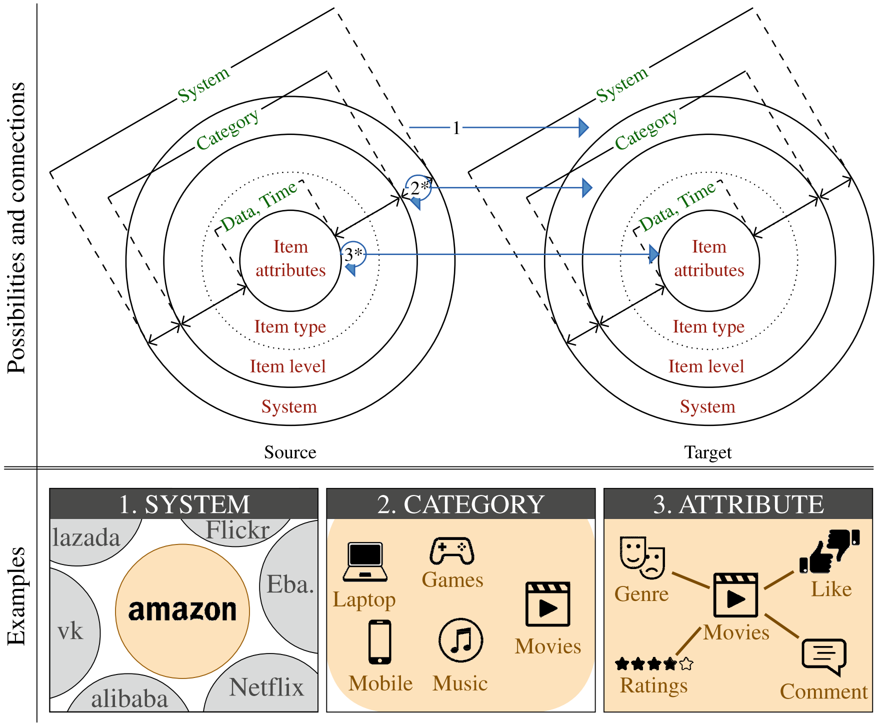


Fig. 3 Domain possibilities, connections and examples

4.1 Domain Selection Criteria

Before dwelling into the explanation of each component it is important to first identify what "domain" means and the relationship between them.

Figure 3 is constructed using "domains" explained by [13, 20]. First, let us focus on the target and domains identified in it. The target consists of four domain types described by [20], shown in red color, and four domain types described by [13], shown in green color.

For "system domain", both agree on same definition, which is *"Items belonging to different systems are considered as belonging to different domains; for example, MovieLens and Netflix are considered as different domains."* [24]. However [13] envelopes item type and item level domain definitions proposed by [5] under "category domain", which is *"Recommender systems items can be grouped with respect to item types as far as they reside inside a single system domain. Hence, each type can be described as a different category. For example, Amazon items having type electronics, clothes, movies, music etc., can be titled as belonging to respective categories"* (Fig. 4).

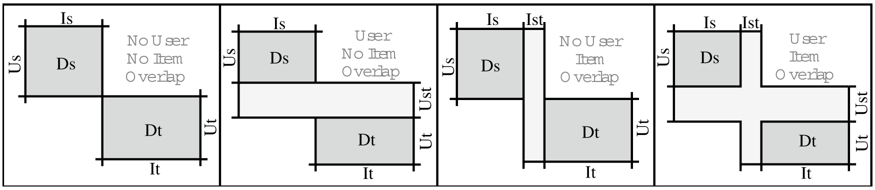


Fig. 4 Overlap scenarios between users and items of the participating domains

Finally, "item attributes domain" definition being broad, encapsulates both times and data domain definitions used by [13], item attribute domain are described as "Two items are considered belonging to different domains if their attribute values are found to be different. For example genre, implicit/explicit data, time etc.".

For different domain, size of the circle represents two things, first, it is proportional to the scope of a rating matrix of a recommender system, e.g., Amazon is a system having a variety of items (Fig. 3, Example 1). Second, each domain having smaller scope can be nested inside the domain having a large scope, e.g., Amazon items can be categorized as movies, music etc., but they reside inside amazon system domain (Fig. 3, Example 2).

This phase focuses on research problem, target and source domain selection criteria, where as, methods of transferring knowledge (shown as blue lines in Fig. 2), will be discussed in next phase i.e., 'Knowledge Transfer'.

4.1.1 Problem Identification

The First task in cross-domain recommender systems research is to identify the problem faced by the target domain. Under cross-domain recommender systems research, problems can be listed under three categories. First, **"knowledge transfer"** is related to problems which are CDRS specific [20]. They focus on different methods of transferring knowledge between the source and the target domain. Second and third i.e., **"group-oriented recommendation"** [4] and **"user-oriented recommendation"** [19] are recommender systems specific problems.

One issue with CDRS specific problems is that usually, the only way to evaluate the performance of proposed methods is to either implement them in combination with group-oriented recommendation algorithms or using user-oriented recommendation algorithms [20]. This is shown as a bright overlapping region containing component 6, 7 and 8 in Fig. 2.

4.1.2 Target Domain Selection

In the field of cross-domain recommender systems, target Domain is assumed weak as compared to the source domain [23], however, empirical truth is that target domain is usually selected as the industry standard, for example, Movielens. This is because the target domain should have enough ratings that can be used for evaluation of the proposed approaches.

Problem identified in first component are simulated using selected target domain, whereas, proposed algorithms utilize data in source domain for improve recommendation in target domain [21].

4.1.3 Source Domain(s) Selection

Once target domain is selected, researchers look for appropriate source domain. Main criteria of selection is that, source domain should contain extra information as compared to target domain. Extra information of source domain, once transferred, improve target recommendation.

One limitation highlighted by different researchers is that CDRS research is lacking sufficient source domain [4]. Also there is a need of methods/techniques which convert data of existing systems such as social media websites, into recommender system compatible data. This will decrease ratio of simulated source domain in the field of CDRS [9].

4.2 Knowledge Transfer

After the identification of the research problem and selection of the target, source domain, knowledge transfer is performed. hence, blue lines shown in Fig. 3 numbered as one, two, three show direction of knowledge transfer.

The first line shows transfer of knowledge from one system domain to another. The second line shows knowledge transfer between categories, it has two variants, first, knowledge from one category can be transferred to another, within the same system domain. Second, knowledge from one category is transferred to another category residing in another system domain. same applies to attribute level knowledge transfer.

This Phase contains four components. "Similarity identification" is related to identification of common features in both source and target domain. "Transformation" is responsible for making features of both domains compatible with the proposed algorithm. Whereas, both "Knowledge aggregation" and "knowledge linkage" are knowledge transfer techniques. "Knowledge linkage" techniques are usually used when no similarity is identified between participating domain.

4.2.1 Similarity Identification

In order to transfer knowledge from source to Target domain some similarity need to exist. Study [7] outlined 4 types of similarities between the source and Target domain. Let us suppose that source domain is represented as Ds having users (Us) and items (Is), on the other hand, Target domain is represented as Dt having users Ut and items It. Reference [7] state that four similarity possibilities exist. First, users and items of both source and Target domain overlap. Second, only users overlap, third, only items overlap and fourth, no overlap exists. By overlap [7] means that uses or items are common between domains.

It is essential to identify similarity because it was found that usually knowledge aggregation (Sect. 4.2) applies to those scenarios where overlap exists. Whereas, for no overlap knowledge linkage (Sect. 4.2) is the only option.

4.2.2 Transformation

Transformation is essential for establishing prediction algorithm's compatibility with source and Target domain. This requirement arises when participating domains have different interaction types, e.g., source domain binary interactions such as likes for the items, whereas, the target domain has numeric interactions such as ratings between 0 and 5. For CDRS There exist two possibilities, **"explicit transformation"** and **"implicit transformation"**. Usually outcome of transformation is a rating metrics, however, some researchers incorporate transformation step into prediction algorithm for avoiding matrix generation.

No transformation is required when source and Target domain ratings have same scale (e.g. [10]). Explicit transformation is required when both domains have numeric ratings but have a different scale (e.g. [16]). Finally, an implicit transformation is required when users provide implicit interactions for items (e.g. [17, 18]). Some example of implicit interactions are like, download, edit, play, click, comment etc., (e.g. [18]). For implicit interactions there exist two main types, first, binary interactions and second, textual feedback. Types of Implicit interactions have grown with the advancement in web 2.0 technologies. Majority researches utilize binary interactions (e.g. [22]),whereas, textual interactions where least found in the selected primary studies (e.g. [8]).

4.2.3 Knowledge Aggregation

Knowledge aggregation is the first type of knowledge transfer technique described by [20], and is described as an aggregation of knowledge from different sources. This technique is further split into three types i.e., Merging user preferences, mediating user preferences and merged recommendation.

- **Merging user preference**: The aggregated knowledge consists of user preferences, e.g., ratings, tags, transaction logs, and click-through data (e.g. [1]).
- **Mediating user preference**: The aggregated knowledge comes from user modeling data exploited by various recommender systems, e.g., user similarities and user neighborhoods (e.g. [3]).

- **combining recommendations**: The aggregated knowledge is composed of single-domain recommendations, e.g., rating estimations and rating probability distributions (e.g. [2]).

4.2.4 Knowledge Linkage

Second technique proposed by [20] was knowledge linkage, where different domain are linked using different models which are trained to find similarities. Features of the participating domain are used for similarity identification and it is not compulsory to have common users or items between participating domain. This technique is further divided into three types i.e., Linking domain, sharing latent features and transferring rating patterns.

- **Linking domains**: Linking domains by a common knowledge, e.g., item attributes, association rules, semantic networks, and inter-domain correlations (e.g. [6]).
- **Sharing latent factors**: The source and target domains are related by means of implicit latent features (e.g. [15]).
- **Transferring rating patterns**: Explicit or implicit rating patterns from source domains are exploited in the target domain (e.g. [16]).

4.3 Recommendation Generation

This phase comprises of three components, prediction algorithm selection, training prediction algorithm and finally, recommendation generation and evaluation. On a broader level, recommendation generation is categorized into either group oriented recommendation or user-oriented recommendation. Component "selecting prediction algorithm" explore both categories in details.

4.3.1 Selecting Prediction Algorithm

Recommendation generation can happen under multiple scenarios such as **group oriented recommendation** [4], **user-oriented recommendation** [19] or both. A group-oriented recommendation is referred to the recommendation which is generated for a group of people, whereas, a user-oriented recommendation is related to recommendation generation for an individual user. Figure 5 shows taxonomy of different approaches under group and user-oriented recommendation. Although researchers are independent to select recommendation scenarios, selection of algorithm will dictate training scheme and evaluation methods.

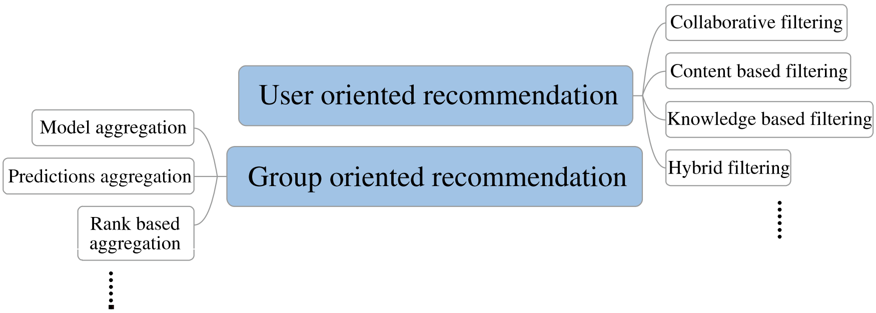


Fig. 5 Basic taxonomy of recommendation schemes

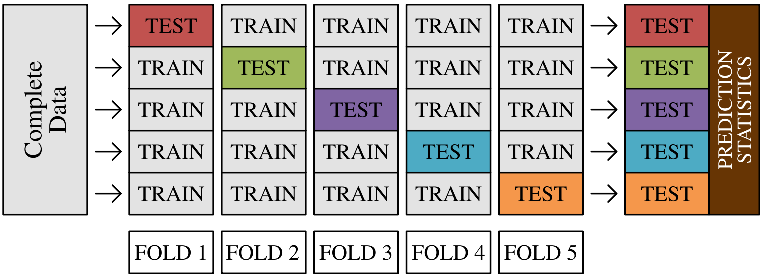


Fig. 6 Cross validation process

4.3.2 Training Prediction Algorithm

Recommender systems rely on a technique called the split validation [11]. In this technique, algorithm is trained on certain percentage of the target domain, whereas, rest of the target domain is used for testing purposes. K-fold validation is also used which repeats split validation K number of times, every time splitting target domain randomly with same ratio. Figure 6 illustrate cross validation steps. This component is same for both recommendation scenarios and this step make sure training process un-biased.

4.3.3 Prediction Generation and Evaluation

Prediction can be either generated for a group of users or for an individual user, therefore, both scenarios are evaluated different. Group-oriented predictions are evaluated using rank-based evaluation methods, such as Spearman rank correlation [4], whereas, user-oriented recommendation algorithms are evaluated using error based evaluation metrics such as mean absolute error (MAE) [11]. Reference[13] has highlighted a variety of evaluation metrics in the response of their third research question.

5 Conclusion and Future Work

This paper presents a generic cross domain recommendation framework which outlines phases and components essential to conduct cross domain recommendation. Proposed framework was based on the analysis of 128 primary studies gathered from two review papers [13, 20]. Although this paper attempts to present example of each component from gathered primary studies, in future we want to utilize proposed generic framework for conducting a systematic literature review in order to identification and classification recently published primary studies. We believe that proposed generic framework can serve as a step-by-step guide for conducting research in cross domain recommender systems.

References

1. Abel, F., Herder, E., Houben, G.J., Henze, N., Krause, D.: Cross-system user modeling and personalization on the social web. User Model. User-Adapt. Interact. **23**(2–3), 169–209 (2013)
2. Berkovsky, S., Kuflik, T., Ricci, F.: Cross-domain mediation in collaborative filtering. In: International Conference on User Modeling, pp. 355–359. Springer, Berlin (2007)
3. Berkovsky, S., Kuflik, T., Ricci, F.: Mediation of user models for enhanced personalization in recommender systems. User Model. User-Adapt. Interact. **18**(3), 245–286 (2008)
4. Cantador, I., Castells, P.: Group recommender systems: new perspectives in the social web. In: Recommender Systems for the Social Web, pp. 139–157. Springer, Berlin (2012)
5. Cantador, I., Fernández-Tobías, I., Berkovsky, S., Cremonesi, P.: Cross-domain recommender systems. In: Recommender Systems Handbook, pp. 919–959. Springer, Berlin (2015)
6. Chung, R., Sundaram, D., Srinivasan, A.: Integrated personal recommender systems. In: Proceedings of the Ninth International Conference on Electronic Commerce, pp. 65–74. ACM, New York (2007)
7. Cremonesi, P., Tripodi, A., Turrin, R.: Cross-domain recommender systems. In: IEEE 11th International Conference on Data Mining Workshops (ICDMW), pp. 496–503. IEEE (2011)
8. Dooms, S., De Pessemier, T., Martens, L.: Movietweetings: a movie rating dataset collected from twitter. In: Workshop on Crowdsourcing and Human Computation for Recommender Systems, CrowdRec at RecSys, vol. 2013, p. 43 (2013)
9. Fernández-Tobías, I., Cantador, I., Kaminskas, M., Ricci, F.: Cross-domain recommender systems: a survey of the state of the art. In: Spanish Conference on Information Retrieval, sn, p. 24 (2012)
10. Gao, S., Luo, H., Chen, D., Li, S., Gallinari, P., Guo, J.: Cross-domain recommendation via cluster-level latent factor model. In: Joint European Conference on Machine Learning and Knowledge Discovery in Databases, pp. 161–176. Springer, Berlin (2013)
11. Gunawardana, A., Shani, G.: Evaluating recommender systems. In: Recommender Systems Handbook, pp. 265–308. Springer, Berlin (2015)
12. Khan, M.M., Ghani, I., Jeong, S.R., Ibrahim, R., Qureshi, K.N.: Facebook's public social interaction utilization to assist recommendation accross system domain. J. Theoret. Appl. Inf. Technol. **88**(3), 392 (2016)
13. Khan, M.M., Ibrahim, R., Ghani, I.: Cross domain recommender systems: a systematic literature review. ACM Comput. Surv. (CSUR) **50**(3), 36 (2017)
14. Khan, M.M., Ibrahim, R., Younas, M., Ghani, I., Jeong, S.R.: Facebook interactions utilization for addressing recommender systems cold start problem across system domain. J. Internet Technol. **19**(3), 861–870 (2018)

15. Koren, Y.: Factorization meets the neighborhood: a multifaceted collaborative filtering model. In: Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 426–434. ACM, New York (2008)
16. Li, B., Yang, Q., Xue, X.: Can movies and books collaborate? cross-domain collaborative filtering for sparsity reduction. IJCAI **9**, 2052–2057 (2009)
17. Lu, Z., Zhong, E., Zhao, L., Xiang, E.W., Pan, W., Yang, Q.: Selective transfer learning for cross domain recommendation. In: Proceedings of the 2013 SIAM International Conference on Data Mining, pp. 641–649. SIAM (2013)
18. Moreno, O., Shapira, B., Rokach, L., Shani, G.: Talmud: transfer learning for multiple domains. In: Proceedings of the 21st ACM International Conference on Information and Knowledge Management, pp. 425–434. ACM, New York (2012)
19. Park, D.H., Kim, H.K., Choi, I.Y., Kim, J.K.: A literature review and classification of recommender systems research. Exp. Syst. Appl. **39**(11), 10,059–10,072 (2012)
20. Ricci, F., Rokach, L., Shapira, B.: Recommender systems: introduction and challenges. In: Recommender Systems Handbook, pp. 1–34. Springer, Berlin (2015)
21. Rosli, A.N., You, T., Ha, I., Chung, K.Y., Jo, G.S.: Alleviating the cold-start problem by incorporating movies facebook pages. Clust. Comput. **18**(1), 187–197 (2015)
22. Sedhain, S., Sanner, S., Xie, L., Kidd, R., Tran, K.N., Christen, P.: Social affinity filtering: Recommendation through fine-grained analysis of user interactions and activities. In: Proceedings of the First ACM Conference on Online Social Networks, pp. 51–62. ACM, New York (2013)
23. Shapira, B., Rokach, L., Freilikhman, S.: Facebook single and cross domain data for recommendation systems. User Model. User-Adapt. Interact. **23**(2–3), 211–247 (2013)
24. Xu, Q., Xiang, E., Yang, Q.: Social-behavior transfer learning for recommendation systems. In: Proceedings of Workshop on Social Web Mining (2011)

Measurement of Service Quality of a Public Transport System, Through Agent-based Simulation Software

Mauro Callejas-Cuervo, Helver A. Valero-Bustos, Andrea C. Alarcón-Aldana and Miroslava Mikušova

Abstract An agent-based modeling software is here presented which simulates the measurement of the quality of the service offered by a collective public transport system, through the evaluation of the variables of comfort and speed. The simulator takes into account the trajectory of a route from the bus terminal to its last stop, pausing at each of the stops in the bus itinerary. The software allows for the configuration of the location of each stop, the speed per segment, the distribution of the generation and attraction of tickets per stop, among others. The output information shows the number of passengers waiting, those who leave, journey time, distance covered, and passengers served. In the trajectory tested, an average of 3.9 was obtained with regard to comfort and a 3.1 with regard to speed, using a scale of 1–5.

Keywords Service quality · Collective public transport · Agent based simulation software · Comfort · Speed

1 Introduction

The planning of a route or a number of routes in an urban collective public transport is a process which is influenced by a number of factors related with the supply and demand which is varying continuously in time and space. For this reason, some adjustments have to be made to their schedule during a period of implementation. This means that the process of scheduling the operation is note exact, but it is subjected to uncertainties. However, not doing so causes series disadvantages in the quality of the service, which are reflected on social and economic negative effects.

Many complex social systems have used simulation techniques to analyze their behavior and plan improvement strategies to obtain the results sought [1–3]. Among

M. Callejas-Cuervo (✉) · H. A. Valero-Bustos · A. C. Alarcón-Aldana
Universidad Pedagógica y Tecnológica de Colombia, Tunja 150001, Colombia
e-mail: mauro.callejas@uptc.edu.co

M. Mikušova
University of Zilina, Žilina, Slovakia

M. Huk et al. (eds.), *Intelligent Information and Database Systems: Recent Developments*, Studies in Computational Intelligence 830,
https://doi.org/10.1007/978-3-030-14132-5_27

those techniques, there are software agents and multi-agent systems, which are programmes that are associated with a behavior and own characteristics which allow for the representation of a really complex system [4–7].

The simulation of the transport systems has allowed for the identification of elements and variables that should be taken into consideration regarding the programming and distribution of resources in different fields, but most of them analyze or research on the optimization of routes in different circumstances [8–10]. It is also necessary to mention that there is little exploration computer wise, suing software agents in the measurement of variables, such as comfort and speed, main components of a complex system [11]. The objective of this work was justified in the use of software agents with own behavior to design, model and simulate the operation of a public transport system in medium-sized cities, using the Bottom-Up, specific for multi-agent systems [12].

This article is structured as follows; first, there is the construction and simulation of the model based on agents that are part of the materials and methods used in the development of the research. Then, there is a record and analysis of the results. Finally, the most relevant conclusions are put forward.

2 Materials and Methods

In the same way that object-oriented methodologies offer guides for the identification of objects and interdependencies between them, there are methodologies for the modelling and design of multi-agent systems. Most of these systems have their grounds in the Unified Modelling Language (UML) [13], with are the basis for the development of software with agents, and among them, the following can be found: BDI, MAS-CommonKADS, GAIA, MASSIVE, INGENIAS. On the other hand, agent-based simulation [14] has guides that show how to decompose a problem and how a design will affect the general functioning of the system. This design can be done from two perspectives: Top-down or Bottom-up. The latter is based on an architecture of reactive agents, which models from particular to general, as it firstly visualizes the system in each one of its parts (unanimated agents); then, rules are assigned to each one of those agents (animated agents); following the union of a society of agents (organization of agents); and finally, there is the analysis of the behavior of said agents or data, a technique that was used as the basis in order to carry out this work.

Agent-based simulation puts forward the idea of developing the following phases: contextualization and analysis of the system to be simulated, modeling of the system, creation of the graphic interphase of the user, creation of scenarios to experiment, simulation of the models with software and collection and analysis of results. Below the first five phases and described and the last two are developed in the results section.

2.1 Contextualization and Analysis of the System to be Simulated

The process of travelling in a means of urban collective public transport starts the moment a person makes the decision of taking this journey from the place where they are to their place of destination. For this, they have to go to the bus stop that offers them the right route of transport. Once there, they have to wait a certain time until the bus arrives and then the person wonders if that is the correct bus. If the answer is negative, they have to continue waiting. If the answer is affirmative, the person now wonders if the bus is the type they want to take. If the answer is yes, the person stops the bus, but if not, a new question arises: “do I have enough time to get to my place of destination?” If the person had the time, they can wait for another bus that adapts more to their taste, but if not, they stop the bus to start their journey. In case that the person makes the signal to stop the bus, but the vehicle does not stop, wait for another vehicle. Otherwise, the passenger analyses if there are seats available. If there are, they get on the vehicle. Once there, the following question arises: are there any seats available? If there are, the person takes a seat, if not, the person remains standing. At every stop, the person wonders if that is their stop. If it is, the person gets off the bus, if not, they remain in their place (if they are seated), but if the person is standing, they analyze if there are seats available to take one. This last process takes place until reaching the destination bus stop. When the route of the passenger finishes, the individual descends from the vehicle and goes to their final destination.

The process described above is carried out by each and every person who uses urban collective public transport according to their age, gender and preferences when choosing transport. They generate the journeys and, in general, they have a reason, a schedule, a departure time from the stop of origin, and arrival time to the destination stop, in addition to an origin and a destination. In relation to these characteristics, there is another important factor for these individuals that is the time and distance from the origin to the bus stop, the time and distance covered, the time and the distance from the bus stop to their final destination.

Another process involved in the transportation of people in an urban collective public transport system is, without doubt, the one carried out by the vehicle. A vehicle starts its process at the bus terminal. The driver analyses if it is time to depart and if he has the authorization to cover the route. In case that he does not fulfill any or none of the characteristics mentioned above, the driver has to wait. If he does meet the criteria, the displacement of the vehicle begins. In turn, the driver has to comply with certain stops of the vehicle, as the compulsory stops (stop signs, turns, traffic lights), and in case he finds any, the vehicle is stopped until he is given the sign or authorization to continue. If a person makes a sign to stop the bus, the bus will only stop if it has enough available seats allowing for the person to make the decision of getting on the bus or not. Once the action is carried out, the vehicle will continue its way. In case of stopping in a bus stop, the vehicle will stop and start its waiting time, when the limit of waiting time is reached or before if the driver decides to, the vehicle continues its way. This process takes place in a similar way until the bus reaches the end of the route assigned.

The prior procedure is carried out by each one of the vehicles, which cover a route and the characteristics of which are: an itinerary that has an origin and a destination of the assigned route and a certain time to cover it, bus stops with pre-defined waiting time for each, passing points, apart from having a fare and a schedule assigned (departure time from the origin point, arrival time at destination).

- Unit of analysis

The representative sample or unit of analysis taken for this research project is given by the selection of a route of congestion and common traffic. The route covers a traffic artery from the city of Tunja. The origin of said route is the bus stop of Barrio Reten Sur and its destination is the bus stop Barrio los Muiscas. The latter is a pole of attraction of journeys at different times of the day, considering that it has an important educational sector (a university, several public and private schools). The route in question is 8.37 km long, which are covered in 35 min (time assigned due to the results generated from a study of routes and implemented by most of the transport companies that operate in the city of Tunja) by a vehicle of urban collective public transport.

2.2 *Modeling of the Agent-Based System*

The implementation of the model was developed with the implementation of the tool AnyLogic. Active classes or agents, among which there are the following: dispatcher, dispatch, periodP, periodT, person, route, site, section, vehicle, journey, with a structure of its own, behavior and exchange of messages. In addition, there is an external way of data collection (DSTPCU, DSTPCUaux, DSTPCUaux1, DSTPCUaux2, DSTPCUaux3), which allow for access to records included in a data base and that store the different scenarios with their parameters associated and used in the simulation.

The structure of an agent in AnyLogic is programmed and managed in the following sections: code, animation, algorithmic function, statechart. The code section describes the source code related to the management of the agent. The behavior of each one of the agents involved in the model is set up in statechart; there, the different states and transitions that the interactive agents may adopt are represented. In the animation section, the action referred to the interactivity of the agent with others within the system is set up.

The state diagrams in AnyLogic are done in the section statechart, in which the behavior of each one of the agents involved in the model is set up. There the different states and transaction that the interactive agents may adopt are represented.

- Logical specification of the agents which interact with the urban collective public transport system.
  The following are the behavior rules of some of the components (agents) of the urban collective public transport system:
  The following are the rules of behavior of some of the components (agents) of the TPCU system.

  – Dispatch
    If the vehicle is programmed, then verify departure time:
    If departure time corresponds to the time scheduled for the vehicle, then allow departure of the vehicle and start the time count of the itinerary.
    If not, wait.
    If the vehicle reaches a bus stop, then, stop the vehicle. If not, continue driving.
    If the vehicle stops at the bus stop, then, wait for the passengers whose destination is that bus stop to descend.
    If a passenger descends then subtract 1 to the volume of passengers that are being transported at that moment in the vehicle and count the statistical information of the passenger.
    If a passenger requests the service, then, add 1 to the volume of passengers that are being transported at that moment in the vehicle.
    If the vehicle stops in the last bus stop, then, store the statistical information of the journey.
    If the vehicle is moving, then, adjust the displacement speed according to the section.
  – Bus stop
    If the present time corresponds to a period, then, start journey according to the arrival rates and other characteristics.
    If not, modify the arrival rates with regard to the period of time of the day.

  – Section
    If you are in a moving vehicle on a certain section, then, verify the speed at which it can move.
    If the speed is higher than that assigned to the section, then, reduce the speed.
    If the time is the same as that of the highest limit corresponding to the period of the section, then, update speed.
  – Journey
    If you are at a bus stop then
    If a dispatch arrives at the bus stop, then, verify that there are seats available. If there are seats available, get on the bus.
    If not, wait.
    If waiting time is greater than the time destined to this purpose, then, abandon the system.
    If not, continue waiting.
    If you are in a dispatch, then:
    If that is the right destination, then, descend from the vehicle.

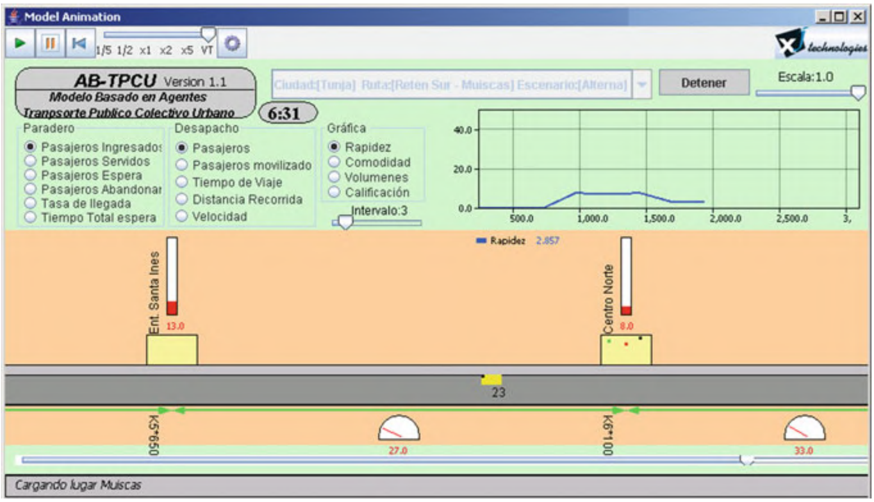


Fig. 1 Graphic user interface of an agent-based simulator

If it is not your destination and you are standing, then, look for an available seat. If there are available seats, sit down.

– Route If a dispatch gets to the destination terminal, then, count the information of the dispatch.

2.3 Creation of the Graphic Interphase of the User

It is in the animation section where the graphic interphase of the user is created and set up, and where the behavior of the agents is shown. In Fig. 1, the different simulation areas.

The section in green includes the elements that allow the configuration of different actions carried out by the simulator. On left hand side of the green area, it is possible to see the title of the application AB-TPCU, as well as its version. Below, there is a set of three selection buttons that read passenger, dispatch and graph. Towards the front of the title of the application, there is information related to the configuration of the simulator, such as name of the city in which the simulation will be applied, in this case "[Tunja]", name of the route "[Reten Sur – Muiscas]" and, finally, the programming of the route of scenario, in this case "[Real]", which is a test programming to conduct the analysis. Beside this information, there is an icon that displays a list of options to select the scenario in which the simulation will take place. Right after it, there is a button that reads "Load data/Stop", which connects the database that contains information about the parameters. This information is related to the distribution of bus stops, sections, itineraries and other elements. Next to this button, there is level bar called "Scale", with which the visual appearance of the route can be augmented or reduced (pink). Below the green area (bottom right), there is a graph that shows the behavior of the urban collective public transport system where the speed, comfort, certain measures of volume and the rating of the service regarding comfort and speed. This is managed through a set of buttons called "Graph". When selecting the type of graph desired, this produces a swap and changes the view to the option selected.

In the same figure, an intermediate area can be observed (pink) where there are the different agents that intervene in the simulation of the urban collective public transport system. They activate when the button "Load data/Stop/Simulate" changes its name or label to "Simular" and it is pressed. The elements related to this area are:

The area represented by pound key 1, yellow rectangle, represents the bus stops and on it, their name can be read, in this case "Centro Norte". The red and blue dots within the yellow rectangle (bus stop) represent the users of the system.

The area represented by pound key 2, in grey, represents the road on which the vehicles circulate and the green rectangle (on the road) represent the buses. The number that appears underneath is the number of passengers the vehicle is transporting at that moment.

The displacement bar (pound key 3) below the pink area allows for the displacement of objects (bus stops, sections and other static elements of the simulation) front or back, allowing for observation of the movement of the vehicle while it passes by the different bus stops of the route. Some elements of the simulator can be set up.

2.4 Creation of Scenarios to Experiment

For the creation of scenarios to experiment, the city of Tunja was selected, more specifically the route Reten Sur – Muiscas, which is one of the most important routes that covers the city from south to north. The information used as basis for this analysis dates from 2003, but it is still valid because the infrastructure has not suffered big changes and the data was provided by reliable sources.

The route has 14 bus stops and an 8.730 m length distributed as follows: (1) Reten Sur (bus stop of origin), (2) Los Hongos, (3) Comboy – Gaseosas, (4) Colegio Rosario, (5) K 9 Con Cll 21, (6) Nieves, (7) Hugolino, (8) Entrada Mesopotamia, (9) UPTC, (10) Santa Inés, (11) Centro Norte, (12) Toyota, (13) Asis, (14) Muiscas (last bus stop). The route has fourteen stops and has a total length of 8,370 m, distributed as follows: (1) Reten Sur (Source stop), (2) Los Hongos, (3) Comboy – Gaseosas, (4)Colegio Rosario, (5) K 9 Con Cll 21, (6) Nieves, (7) Hugolino, (8) Entrada Mesopotamia, (9) UPTC, (10) Santa Inés, (11) Centro Norte, (12) Toyota, (13) Asis, (14) Muiscas (Destination stop).

- Real scenario: the programme of the route for a working day in the city of Tunja was taken into consideration. It was a Friday, given that on this day the social and economic activities of the city increase. The information of the base parameters of the simulation is observed in Table 1.

Table 1 Simulation scenario: real data

Parameter	Value parameter
Route	Reten Sur – Muiscas
Day	Friday
Time	6 a.m. 10 p.m.
Number of vehicles	27 Buses
Dispatches per day	65 Buses/Day
Maximum vehicular capacity	2500 journeys
Time interval between dispatches	15 min

- Alternate scenario: it starts from the values shown in Table 1, but it decreases the time between shipments (sending vehicles to provide the TPCU service) in peak periods, in hours between 7:00 a.m. and 9:00 a.m. (dispatches every 5 min) and in the valley period of 6:00 a.m. at 7:00 a.m. and from 9:00 a.m. at 11 a.m. the same times are maintained (dispatches every 15 min). In this scenario, the same automotive park of the real programming is considered.

2.5 Simulation of the Models with Software

The models are simulated using AnyLogic, taking a real and alternative scenario as its basis. To analyse the behavior of the variables associated with agent-based simulation a range from rush hour is taken into consideration, 7:24 a.m. and 8:24 a.m. It is important to take into account that in the model, speed, comfort and rating are recorded in the moment that each one of the passengers finish their journey.

3 Results

Below there is a description of the behavior of the speed and comfort variables, taking as a reference both, the real or base and the alternative scenarios. There is also an analysis of the quality evaluation.

3.1 Preliminary Results

- Speed variable. Analysing Fig. 2a, b, it can be concluded that the average difference of the speed between real and alternative scenarios is approximately of 3 km per hour (3km/h), which in terms of time, for the users, it implies saving 3 min of the

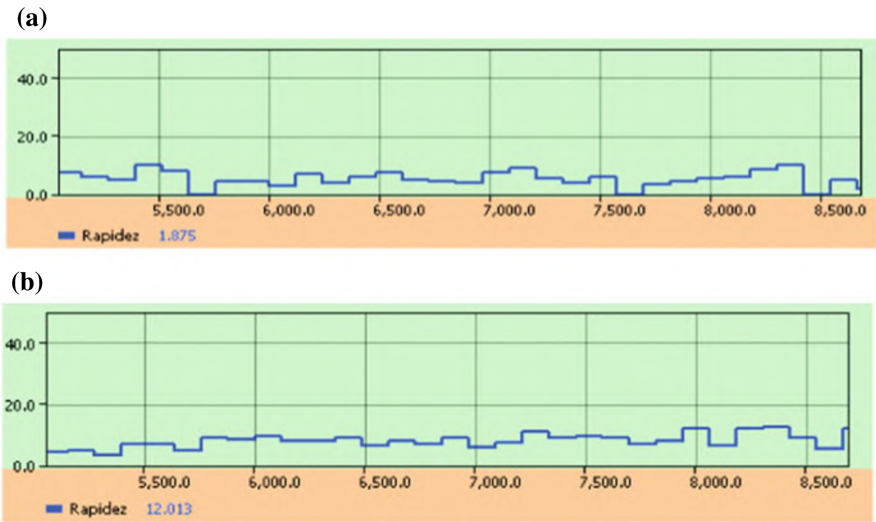


Fig. 2 Behavior of the speed variable. **a** Speed variable: real scenario. **b** Speed variable: alterative scenario

journey. The time saved is only related to the waiting time of the users at the bus stop, given that the displacement speed of the vehicles is relatively constant. In addition, there is some improvement in the behavior of speed in the alternative scenario.

- Comfort variable. Figure 3a, b represent the behavior of the comfort variable, where it can be seen that there is a greater degree of comfort in the alternative scenario, given that there is an increase in the frequency of dispatches and the users will have more chances to sit during the journey.

  Quality rating. The perception of the user about the quality of the service, regarding comfort and speed is presented in a scale of 0.0 to 10.0, where 0.0 is a deficient quality and 10.0 is an excellent quality. Figure 4 shows this behavior, where the alternative scenario shows a more uniform rating than the real scenario, reflecting that the quality of the service would improve with the conditions set up in the scenario proposed. In addition, it shows that the rating for comfort in the real scenario is approximately 3.0 with a high deviation whereas the rating for the alternative scenario is approximately 5.0, with a low deviation. This same behavior is shown in the speed rating where it is 1.5 for the real scenario and 2.5 for the alternative.

3.2 Result Analysis

The evaluation of the model reflects the importance of an adjustment in the operation programming of the public transport system to each period of the day (rush hour and normal)

(a)

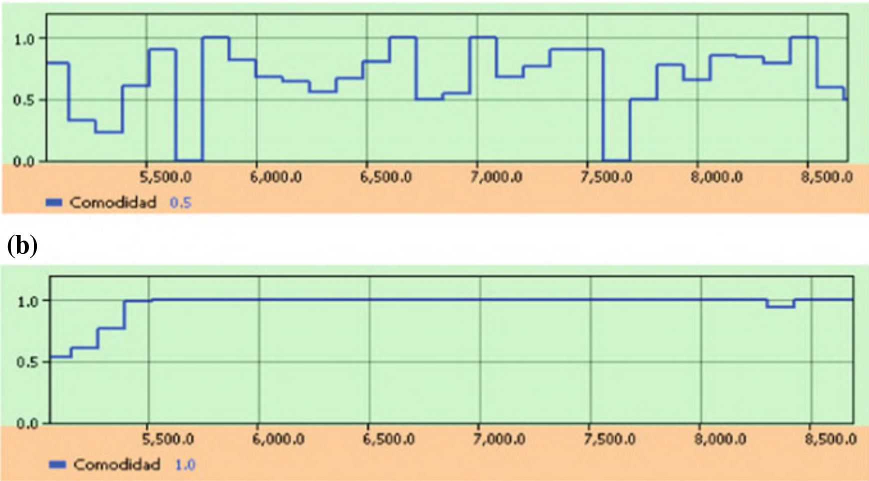


Fig. 3 Behavior of the comfort variable. **a** Comfort variable: real scenario. **b** Comfort variable: alternative scenario

(a)

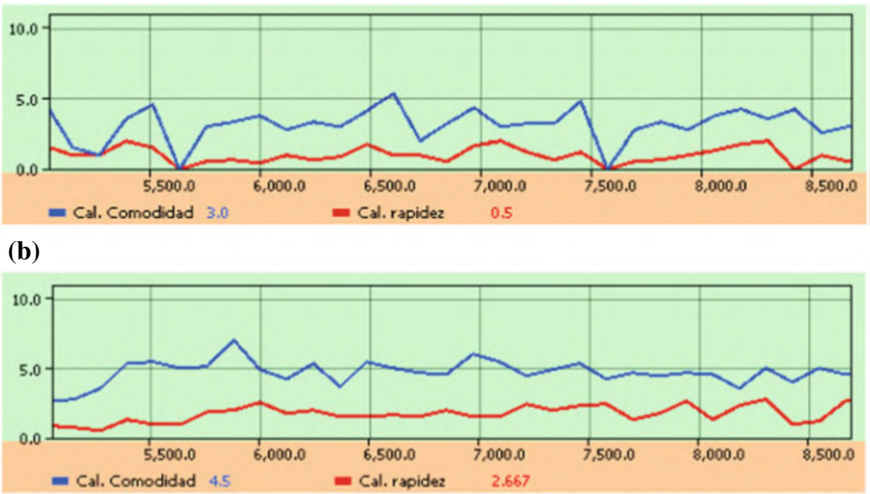


Fig. 4 Rating of the service. **a** Quality: real scenario. **b** Quality: alternative scenario

The average results (condensing the information in a day by the transport system, in the real or base situation) for each one of the variables were:

- Level of service rated for comfort: 3.8.
- Level of service rated for speed: 2.5.
- Final rating of the service: 3.15.

For this, it can be said that the service, according to the system, was rated as AVERAGE.

The average results in the agent-based simulation (condensing the information generated in one day in the system, alternative scenario) for each of the variables if analysis with regard to the factors were:

- Level of service rated for comfort: 3.90.
- Level of service rated for speed: 3.10.
- Final rating of the service: 3.50.

It can be concluded that the quality of the service, according to the system, is GOOD.

In Table 2, the results obtained after the simulation of other scenarios with some variations in the values of time regarding the departure of vehicles.

The results shown in the previous table were taken into consideration with the purpose of validating and testing the models developed. As it can be observed, several

Table 2 Results of the different executions of the model

Schedule	Test	Comfort rating	Speed rating	Number of daily journeys
Real	0	3.75	2.51	4024
Real	1	3.70	2.48	4080
Real	2	3.73	2.49	4158
Alternative	0	3.77	2.47	4068
Alternative	1	3.76	2.46	4038
Alternative	2	3.73	2.46	4131
Alternative 1	0	3.70	2.53	4021
Alternative 1	1	3.69	2.53	4013
Alternative 1	2	3.70	2.54	4026
Alternative 2	0	3.85	2.51	4051
Alternative 2	1	3.84	2.47	4068
Alternative 2	2	3.87	2.51	4009
Alternative 3	0	3.71	2.48	4089
Alternative 3	1	3.71	2.50	4157
Alternative 3	2	3.70	2.46	4056

tests were carried out in each one of the scenarios in order to have an average of the ratings and evaluate the behavior with the different variations in the dispatch times (time frequency with which buses leave in order to cover the full route).

4 Conclusions

The importance of an assertive operational planning lies in managing to give a service with acceptable quality for the users and, thus, avoid that they change to other systems of transport with a higher economic, social and environmental cost. For this reason, it is crucial to carry out a follow-up, through the application of the model that has been developed, to routes that present problems with journey circulation, with the aim of evaluating the behavior of the system and making a decision regarding an adjustment in the programming that leads to an improvement in the service for this sector of the population and, as a consequence, benefit the companies that provide the service. The modelling system reached a high level of precision regarding details, as was the case of managing to follow up a vehicle along its whole route and register its speed and capacity. On the other hand, this model allowed for the analysis of the supply and demand behavior of the journeys in different bus stops and for showing how certain schedule affects them in different ways at different times of the day.

It is absolutely feasible to use the simulation models presented in this work in other cities with similar characteristics. The success of the results would be related to the quality of the compilation, classification and processing of data necessary to be integrated to the basic data structure required by the applications developed here.

The construction of models of this type together with the implementation of techniques for the collection of permanent data of the system will allow for the constant monitoring and evaluation of the behavior of the system. Thus, it will provide the tools to make better and prompt decisions in relation with the schedulling of the operation of a collective public transport system, apart from allowing a first approximation to intelligent transport systems.

References

1. Davidsson, P., Klügl, F., Verhagen, H.: Simulation of complex systems. In: Magnani, L., Bertolotti, T. (eds.) Springer Handbook of Model-Based Science. Springer Handbooks, pp. 783–797. Springer, Cham (2017)
2. Adelt, F., Weyer, J., Hoffmann, S., Ihrig, A.: Simulation of the governance of complex systems (SimCo): basic concepts and experiments on urban transportation. J. Artif. Soc. Soc. Simul. **21**(2), 1–2 (2018)
3. Kelley, B.N., Williams, R.A., Holland, J.L., Schnarr, O.C., Allen B.D.: A persistent simulation environment for autonomous systems. In: Aviation Technology, Integration, and Operations Conference. Atlanta, Georgia (2018)

4. Carole, A., Patrick, T., Julie, D., Benoit, G.: BDI vs FSM agents in social simulations for raising awareness in disasters: a case study in Melbourne bushfires. Int. J. Inf. Syst. Crisis Response Manag. **9**(1), 27–44 (2017)
5. Bourgais, M., Taillandier, P., Vercouter, L.: Enhancing the behavior of agents in social simulations with emotions and social relations. In: Dimuro, G., Antunes, L. (eds,) Multi-Agent Based Simulation XVIII. MABS 2017. Lecture Notes in Computer Science, vol. 10798 (2018)
6. Mozhgan, K.C., Paul, R.: Simulating heterogeneous behaviors in complex systems on GPUs. Simul. Model. Pract. Theory **83**, 3–17 (2018)
7. Hulme, A., Thompson, J., Nielsen, R.O., Read, G.J.M., Salmon, P.M.: Towards a complex systems approach in sports injury research: simulating running-related injury development with agent-based modelling. Br. J. Sports Med. 1–11 (2018)
8. Wu, N., Wenbo, L., Hailei, W., Chengxin, X., Dan, G.: Capacity-constrained contraflow adaption for lane reconfiguration in evacuation planning. J. Adv. Transp. 14pp. (2018)
9. Eric, N.B., Charles, A.A., Kwame, K.O.: Volume warrants for major and minor roads left turning traffic lanes at unsignalized T-intersections: a case study using VISSIM modelling. J. Traffic Transp. Eng. **5**(5), 417–428 (2018)
10. Novikov, A., Zyryanov, V., Feofilova, A.: Dynamic traffic re-routing as a method of reducing the congestion level of road network elements. J. Appl. Eng. Sci. **16**(1), 70–74 (2018)
11. Nigel, G., Troitzsch, K.G.: Simulation for the Social Scientist, 1^a edn, pp. 67–69. Open University Press, New York (2005)
12. Valentino, C., Aram, G., Kristina, L.: Top-down vs bottom-up methodologies in multi-agent system design. Auton. Robot. **24**(3), 303–313 (2008)
13. Rumbaugh, J., Jacobson, I., Booch, G.: The Unified Modeling Language Reference Manual, 1st edn. Addison-Wesley Professional, Massachusetts (1999)
14. Fikar, C., Hirsch, P., Nolz, P.C.: Cent. Eur. J. Oper. Res. **26**(2), 423–442 (2018)

Detection of Drivers Plate at Smart Driver's Score Application Controlled by Voice Commands

Ondrej Vondra, Jan Dvorak, Ondrej Krejcar and Peter Brida

Abstract These days an availability of mobile technologies and smart devices is still growing and thus a quantity of applications developed for such devices increases. These applications contain a huge number of possibilities, as for example utilization of a camera to get image data and consequent deduction of information out of these data. Also identification of voice commands used to control applications by simple instructions has a large utilization. Even this paper deals with connecting these two possibilities, concretely by rating drivers on the basis of voice commands. The application uses reading image data to gain licence plate numbers of vehicles. For licence plate numbers recognition this application uses openALPR API, which ensures a high accuracy even on less quality images and weak devices. The application created in this work is developed primarily for Android platform. Therefore, the intention of the paper is to create such application.

Keywords Android · ALPR · Driver · Rating · OCR · Voice recognition

1 Introduction

The area of mobile devices is a very quickly developing field with a high availability. Nowadays mobile devices simplify life of most people; either in the scope of entertainment or work, mobile devices are always at hand [1]. One of the main reasons

O. Vondra · J. Dvorak · O. Krejcar
Faculty of Informatics and Management, Center for Basic and Applied Research, University of Hradec Kralove, Rokitanskeho 62, Hradec Kralove 500 03, Czech Republic
e-mail: ondrej.vondra@uhk.cz

J. Dvorak
e-mail: jan.dvorak@uhk.cz

O. Krejcar
e-mail: ondrej.krejcar@uhk.cz

P. Brida (✉)
Faculty of Electrical Engineering and Information Technology, University of Zilina, Univerzitna 1, 010 26 Zilina, Slovakia
e-mail: peter.brida@fel.uniza.sk

M. Huk et al. (eds.), *Intelligent Information and Database Systems: Recent Developments*, Studies in Computational Intelligence 830,
https://doi.org/10.1007/978-3-030-14132-5_28

of this development is their usability, particularly in terms of various applications. Modern applications are being created with as little user interaction as possible—the more work the application can do for the user, the better [2]. Such mobile application brings benefits also to drivers, who can't work with the mobile phone while driving the way that some applications require [3]. An example of applications for drivers can be various Dash video cameras, which only record the journey or process image in some way. One of the more complicated applications is DriveSafe [4]. This application rates a driver pursuant to his behaviour.

There are many applications that use a mobile video camera, but in most of the cases it's the kind of an application that only records an image, but doesn't put any added value to it. This is unfortunate, because it's possible to gain the licence plate number of a vehicle moving in front of the user's vehicle using such application by a simple image capture with OCR utilization [5, 6]. This way of using OCR (Optical Character Recognition) is called ALPR (Automatic Licence Plate Recognition [7]) and it is clearly explained in the article [8, 9]. A very useful element of these applications is a voice control. It enables the driver to control the application by simple commands without losing control of the driving. The technology that allows acquiring voice commands is called ASR (Automatic Speech Recognition) and is being used mostly in connection with automotive industry [10]. This connection is apparent mainly in the navigation systems, which use the voice commands to change a rout, and the like [11].

By connecting the aforementioned approaches and technologies it can be created a very interesting application, which also happens to be the main intention of this work. This application will serve to rate drivers by a voice evaluation as in the school. This evaluation will take place after reading a licence plate of the vehicle moving in front of the evaluating driver. To read the licence plate from an image there can be used many ways and libraries of third parties [12, 13]. The easiest and the most reliable solution is using openALPR API, which is able to return licence plate with vehicle information on the basis of image data. This way there can be achieved surprisingly good results even in adverse conditions. If the read licence plate number is found in the database, the user gets a rate made as an average value of all the evaluations. Otherwise, the application only asks the user to record his own evaluation. Each user can rate one licence plate number only once, but has a possibility to change the evaluation later on. The application will scan the image and gain the information about the vehicle moving in front of the user's vehicle with a certain frequency. For a correct performance of the application it is supposed a large community of drivers, because there's a huge number of vehicles and it would take a very long time for one driver to collect information and evaluation for a sufficient number of vehicles.

The application will be created for Android system [14], because it's the most widespread mobile system and thus it can ensure a huge community of drivers, which happens to be the main factor of this application's success.

2 Problem Definition

As it was mentioned, the main intention of this work is to design and create an application concerning with rating drivers on the basis of user defined or default commands. Hand in hand with this evaluation it will be necessary to identify concrete vehicles by a unique identification, which is their licence plate number. Thanks to this identification the application is able to say in advance, who is the driver that drives in front of us, and the user gets the collected evaluation about the vehicle if it's available.

The main precondition of the application's proper function is a very extended user community of Android mobile devices [15]. Thanks to these users there will be data collected in the form of rating the individual vehicles and consequently on their basis there will be created overall evaluations of the individual vehicles.

This problematic isn't very frequent and developers deal with it only seldom, nevertheless, there exist several solutions similar to the one described by this work. Most of the solutions resolve monitoring of the drivers' behaviour, either the ones moving in front of us or directly us.

A very interesting approach is to analyse driver's behaviour on the basis of common actions and the way he realizes them [16], which deals with the abovementioned analysing the driver's behaviour, even from the moment of opening the car door. These data are being recorded and there's predicted a result driver's behaviour according to them. There was verified a hypothesis, that circumstances and the driver's behaviour can predict one's driving style even before starting a car.

Authors [12] described a mobile application for Android platform utilizing Teseract OCR and a neural network to acquire licence plate numbers by an automatic image recognition using a mobile device. Such gained data can be easily manipulated and they are saved for a later use.

An article [17] describes a portable ALPR implemented on an Android mobile device. The ALPR implementation on a mobile device is markedly difficult, because the mobile devices are not made to process such complicated calculations, therefore the algorithm had to be optimized and simplified very well [18]. There was implemented a technology of multiphase image processing, which first modifies the raw image using contrast and then the image is segmented and there are being acquired the characters of the licence plate by Teseract OCR.

The application introduced in the article [4] was developed only for Apple devices. It's DriveSafe application, which is able to detect patterns of behaviour on the basis of intern sensors and to inform the user about it properly. The application evaluates only events induced by the owner of the device.

This chapter showed several solutions that are already available. However, any solution that would allow the global cooperation of drivers' community is not found. Very interesting is an approach handled by the article [12], which enables to retrieve information using a mobile device and consequently to save them and manipulate with them as it deems appropriate.

Table 1 Solutions comparison

	Platform	Problem solution
[4]	N/A	1
[5]	Android	3
[6]	Android	2
[1]	iOS	3
New solution	Android	5

In Table 1 there can be clearly seen a comparison of the presented articles with the required new solution. One of the elements of evaluation is the problem solution itself. This solution was evaluated within a scale from 1 to 5, while the new custom made solution got the highest rating of 5 points. The other solutions were evaluated according to the available functions and the method of their problem solution, while there was verified their usability and applicability within solving a given problem. A next examined element was the development platform. This criterion was not essential, but it was worthwhile to pinpoint the details and the differences.

The solution resolved by Bergasa et al. [4] may appear sufficient, but unfortunately it supports only scanning the behaviour of the driver, who uses this application and it doesn't enable any communication among different devices, because it works only locally. Moreover, this is the only solution out of the others created just for the iOS platform. A next possible solution seems to be the solution described by Do et al. [12]. This solution only scans the licence plate numbers using ALPR and saves them for a further manipulation. While using this method it would be possible to connect this already existing solution with the newly created one. Nevertheless, this solution would necessarily need a development of an additional solution and as a result the combination of this solution with the new solution could not fulfil the requirements, either functional or performance.

This chapter has shown many possible solutions, and although they are very interesting and instructive, they do not meet the required solution functionalities. This solution will have to be created newly. None of the shown solutions actually doesn't enable saving into a global database with a possibility to retrieve proper and global drivers' evaluations, therefore it is necessary to create a solution that will allow such functionalities.

3 New Solution

The previous chapter presents several possibilities of solution for drivers' rating. This chapter describes an approach selected for the implementation of our solution.

A foundation of the application is a specific kind of image recognition called ALPR, which uses image segmentation to detect licence plate numbers of vehicles and transforms these licence plates to an easily workable text. This mean is clearly explained by Vékony [8]. The simplest way to implement this kind of image recognition is an implementation using API calls. This method has its advantages and disadvantages. One of its main benefits is a lower requirement on performance of the device itself and an ease of deployment. The disadvantage is a necessity of the internet connection, but that is not a major problem nowadays. As an API provider it was chosen openALPR cloud API, which offers a free API option with a big number of calls. The advantage of this provider is a possibility of returning the information from the server. Besides licence plate numbers with percentage results this information includes a colour, car producer and type, which enrich the application with graphic elements and overall user friendliness.

A next important part of the system is divided into two modules, whereas both of them are very important.

The first of these modules is an information output for the user, which will be solved by voice commands using an integrated library TextToSpeech by Google. More detailes about the ASR approach can be gained from article by López-Espejo et al. [11]. This library allows a direct reading of the text data from the application. In the default setting the application will use drivers' rating as in school (from 1 to 5). But the user will have a possibility to change his voice evaluation at his discretion.

The second of these modules is an information input from the user. This module differs from the previous one by its complexity, because retrieving the sound data and their consequent transformation into the text form is quite more complicated operation than the reverse version. Fortunately, even in this case Google Company offers an integrated library for speech-to-text conversion [19]. In this case there's implemented an algorithm for detecting the most similar results of the text input. Pursuant to the percentage results there will be selected the most probable one and the rating will be saved.

The application will be available just for the signed users and their data will be saved either locally or on a central server. For saving the data on the central server it's created REST API, which operates the data on the server [20]. This API will be processed by Microsoft ASP.NET server with a MS SQL database [21–23], which will maintain the data of all the users for their easy access. The application will dispose even with its own "local" database. In this database there will be saved the data needed for the application run and the data, which are not synchronized with the server.

The application will work in several functional modes. These modes are the following: Local, Server and Combination of these two. While using the local mode it will not be asking server for rating of the vehicle in front of us and the data will be saved just locally into the device's memory. In contrary to that the server mode won't save the data halfway in the local database, but in each case of loading or saving the data it will reach directly to the server. The combination mode will first find out, whether there is an information about a given vehicle in the local database and if not, it will ask the server rating.

Generally, the application is divided into two main modes that separate two main functionalities. These modes are Mode of viewing and management of the gained

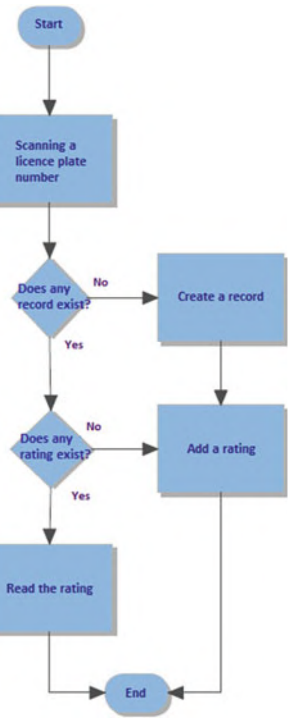


Fig. 1 Diagram of adding a licence plate number

ratings with a possibility of small edits and an application functionality setting. The second mode is Mode of image scanning called „Black box“. In this primary implementation it serves just to capture an image to identify the licence plate number of the car moving in front of us, but later on it can serve as full-fledged „Black box“ for saving a journey record.

In the application there's enabled a manual evaluation input. This evaluation is then realized by taking a picture by a camera or loading an already existing picture of the vehicle from the gallery. Next possibility is to enter all the inputs manually.

In Fig. 1 you can clearly see the application's functionality in the mode of reading the license plate number.

4 Implementation

The application is developed for Android system, concretely for version 5.0 and higher [15, 24]. From this version higher it can be easily used implemented CameraView and library Google TextToSpeech [25]. There are many reasons for choosing this platform; mainly that it is widespread among a general public, which is important for a respectable application function.

In the first development phase it was necessary to define a data structure of the application. Because the application depends on its proper server with API, primary it was needed to create and properly test the functionality of this API. Concretely it's REST API providing and receiving data in JSON format [26]. On the side of the server these data are being stored in a database. As it was aforementioned, for the database technology it was used Microsoft SQL server.

A next phase was processing data objects inside the application. These objects are Plate, Score and User and they represent a key core of the application; they are an essential basis for functioning of the data access and of the whole application. Since the application uses even a local storage, it was necessary to create a local database for saving the needed information. This database was created using automated mapping by ORM framework DBFlow, which automatically maintains the database structure by virtue of the annotated classes. The local database differs from the server one in small details, however, they are just requisite additional information for maintaining the data consistence.

Accordingly, when there are the data accesses created both on the server and on the local side and their functionality is verified, this is the time for a further development phase, which is preparation of a logic layout and an application functionality itself. The application is based on an elementary layout, which fulfils modern standards though; it contains a left scrolling menu and fragment screens. The layout was toned into orange colours and the whole application got the same touch.

An important part of the application is login and registration of the users, which ensures the whole system function. Logging in and registration is the only moment, when the application necessarily requests the connection, because it's needed to identify the user.

Once the application logic structure was designed, there were prepared service classes and methods operating the data inputs and outputs into the application. That is in relation with one of the last parts of the application, which contains an integration of needed libraries and third parties' API. Primarily it's openALPR, which provides acquiring vehicle information from the obtained image data, because this way the data are being acquired in particular. Next it was implemented the user input and output of voice information using Google Speech API, which not only receives the voice commands, but it is also able to read them.

Whereas the application was developed only in a basic mode until this moment, now the time has come for implementation of additional settings that enrich the application with the enhanced possibilities. These settings are described in the following text.

The application contains several setting possibilities; these settings extensively influence behaviour of the whole application. These possibilities include a setting of a mode for licence plate numbers comparison, synchronization mode, which determines intervals for synchronization with the global database. Next there is a scanning interval within „Black box" mode, which determines the frequency of taking pictures and consequently calling the API. The last setting option is a setting of special commands for evaluation; these commands are being used both for entering and for reading these user ratings.

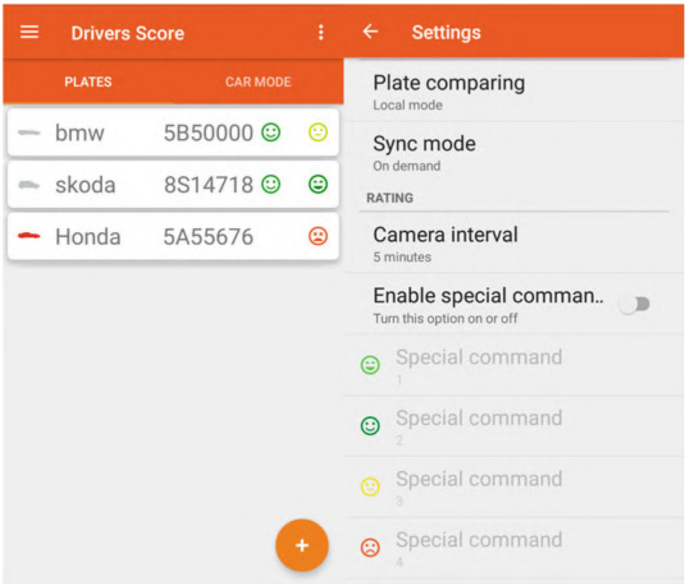


Fig. 2 Screenshots (main screen, settings screen)

In Fig. 2 you can see the screenshots showing the application appearance. In the left screen there's visualized the main screen in the record management mode. In the right screen you can see the settings screen displaying all the application's settings.

The last development phase was implementation of a simultaneous vehicles scanning, which was realized in a form of an image output; in the Fig. 2 you can see a „Car mode" tab in the left screen. In this tab it is displayed the CameraView, which visualizes the camera image on the device's display. The image scanning and gaining information about the license plate numbers is being realized with a given frequency in order to increase the efficiency and improve the performance of the application.

The application remembers the last scanned vehicle and while detecting this vehicle it doesn't realize any user interaction.

5 Testing of Developed Application

In this chapter there are described test scenarios allowing evaluation of the application's functions.

The main test criteria are speed and accuracy of the application [2, 27]. Speed test is provided by measuring speed of comparing and searching the rating of the scanned licence plate number [12, 13]. This test is being realized by measuring speed of executing the performed code with various settings of the searching modes. The types of modes are the Local mode, which compares the data just in the local database,

Table 2 Speed tests

Test number	Local mode (ms)	Server mode (ms)	Combination mode (ms)
1	3	47	55
2	4	43	5
3	3	59	6
4	5	56	4
5	3	37	4
6	2	116	4
7	3	91	4
8	3	51	2
9	4	36	6
10	4	95	5
Average	3.4	63.1	9.5

the Server mode, which downloads and compares the data from the global server and the Combination mode, which combines both of the mentioned modes. The test is being realised by repetitive scanning of the same car image from the gallery using a manual input. There were inserted several time records into the resource code, which document the start and the end of executing the search and the comparison.

The accuracy test is divided into two test phases. The first phase is testing of a success rate of identifying the licence plate number itself. This test depends primarily on the openALPR API, which provides the identified plates. The test is realized by consequent uploading car pictures with various image quality into the application. The second accuracy test phase is testing the speech-to-text conversion. The test is being realized initially while using the special commands and then while not using them. Again there are being tested more cases, when the ambient noise gradually increases.

The results of both of the tests are clearly documented in the following table.

In Table 2 there is clearly displayed the result of testing the speed of searching and comparing the cars' licence plate numbers. There were realized 10 measurements with the same car image [28, 29]. The unit of measurement is millisecond. The results show apparently that the local mode was the fastest; its average speed was 3.4 ms. The second fastest seems to be the combination of the local and server variant, but unfortunately, this result can't be considered as relevant, because for the first time the search took place in both of the databases, but within all the other attempts the information was already saved in the local storage, therefore it wasn't necessary to search on the server. Accordingly, the tests obviously prove that the fastest searching is in the local database, but there can't be guaranteed the most actual data.

The test of accuracy of reading the licence plate number from the image is being realized in terms of quality of the provided image. The test results are displayed in Table 3. These results prove that the results get accurate only when using devices of a higher quality, where the results are very satisfactory. Thus, the total accuracy

Table 3 Accuracy test (image)

Test number	Image quality	Result
1	400×300 px	Not identified
2	640×480 px	Not identified
3	800×600 px	Identified with an error
4	1024×768 px	Identified with an error
5	1600×1200 px	Identified
6	2272×1704 px	Identified
7	2816×2112 px	Identified
8	3264×2448 px	Identified
9	3648×2736 px	Identified
10	4096×3072 px	Identified
Total		6/10

Table 4 Accuracy test (Voice)

Numerical	Ambient noise	Result
1	1	Correct
2	3	Correct
3	5	Correct
4	7	Correct
5	10	Impossible
Special	Total	4/5
1	1	Correct
2	3	Correct
3	5	Correct
4	7	With an error
5	10	Impossible
Total		3/5
Total from both		7/10

happens to be 6/10. But there are included even devices with really low image quality, so we can say that the results are more than sufficient.

The last test is the test of accuracy of reading the user voice input. There was tested a scenario of entering an evaluation “four” for a numerical input and a word “Terrible” for special commands. Both of the scenarios were tested in a various ambient noise simulated by a radio with volume set from one to ten. The test results in Table 4 indicate that in a high level of noise the application can’t identify the command. However, the numerical entry is more accurate, because it uses easier verbal commands and the user is able to define very similar commands.

The realized tests are rather of an indicative nature, because they are realized on a small sample. However, we can say that the application is fast and accurate in most of the cases.

6 Conclusions

The intention of this work was to create an application for rating of the drivers' behaviour based on the voice commands. The application was created for the popular Android platform. We can find two main functions in the application; gaining a rating about the car moving in front of us and entering our rating. The application test was realized by basic tests of speed and accuracy. Both of the tests resulted relatively well and revealed that the application needs a higher quality device with a very good camera and a well shielded microphone.

The application depends a lot on a community cooperation, because it gathers ratings of the other users and creates the average ratings out of them. Thanks to the application drivers will get enough information, which can help to prevent a significant number of accidents.

Acknowledgements This work was partially supported by SPEV project "Smart Solutions in Ubiquitous Computing Environments", 2018, University of Hradec Kralove, Faculty of Informatics and Management, Czech Republic and by project of the Slovak VEGA grant agency, Project No. 1/0626/19.

References

1. Rexha, B., Shala, G., Xhafa, V.: Increasing trust worthiness of face authentication in mobile devices by modeling gesture behavior and location using neural networks. Futur. Intern. **10**(2) (2018)
2. Bartuskova, A., Krejcar, O., Sabbah, T., Selamat, A.: Website speed testing analysis using speedtesting model. J. Teknol. **78**(12–3), 121–134 (2016)
3. Brida, P., Machaj, J., Benikovsky, J., Brida, P., Machaj, J., Benikovsky, J.: A modular localization system as a positioning service for road transport. Sensors **14**(11), 20274–20296 (2014)
4. Bergasa, L.M., Almeria, D., Almazan, J., Yebes, J.J., Arroyo, R.: DriveSafe: an app for alerting inattentive drivers and scoring driving behaviors. In: Presented at the IEEE Intelligent Vehicles Symposium, Proceedings, pp. 240–245 (2014)
5. Behan, M., Krejcar, O.: Vision of smart home point solution as sustainable intelligent house concept. In: Presented at the IFAC Proceedings Volumes (IFAC-PapersOnline), vol. 12, pp. 383–387 (2013)
6. Alpar, O., Krejcar, O.: Virtual road condition prediction through license plates in 3D simulation. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 9875, pp. 269–278. LNCS (2016)
7. Rizvi, S.T.H., Patti, D., Björklund, T., Cabodi, G., Francini, G.: Deep classifiers-based license plate detection, localization and recognition on GPU-powered mobile platform. Futur. Intern. **9**(4) (2017)

8. Vékony, A.: Speech recognition challenges in the car navigation industry. In: Ronzhin, A., Potapova, R., Németh, G. (eds.) Speech and Computer, vol. 9811, pp. 26–40. Springer International Publishing, Cham (2016)
9. Mistry, J., Misraa, A.K., Agarwal, M., Vyas, A., Chudasama, V.M., Upla, K.P.: An automatic detection of helmeted and non-helmeted motorcyclist with license plate extraction using convolutional neural network. In: Presented at the Proceedings of the 7th International Conference on Image Processing Theory, Tools and Applications, IPTA 2017, vol. 2018-January, pp. 1–6 (2018)
10. Maresova, P., et al.: Consequences of Industry 4.0 in business and economics. Economies **6**(3), 46 (2018)
11. López-Espejo, I., Peinado, A.M., Gomez, A.M., Gonzalez, J.A.: Dual-channel spectral weighting for robust speech recognition in mobile devices. Digit. Signal Proc. **75**, 13–24 (2018)
12. Do, H.N., Vo, M.-T., Vuong, B.Q., Pham, H.T., Nguyen, A.H., Luong, H.Q.: Automatic license plate recognition using mobile device. In: Presented at the International Conference on Advanced Technologies for Communications, pp. 268–271 (2016)
13. Alpar, O., Krejcar, O.: Virtual road condition prediction through license plates in 3D simulation. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 9875, pp. 269–278. LNCS (2016)
14. Racko, J., Brida, P., Machaj, J., Krejcar, O.: Path estimation from smartphone sensors. Stud. Comput. Intell. **769**, 435–443 (2018)
15. Marek, T., Krejcar, O., Selamat, A.: Possibilities for development and use of 3D applications on the android platform. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 9622, pp. 519–529 (2016)
16. Sysoev, M., Kos, A., Guna, J., Pogačnik, M.: Estimation of the driving style based on the users' activity and environment influence. Sensors **17**(10), 2404 (2017)
17. Mutholib, A., Gunawan, T.S., Chebil, J., Kartiwi, M.: Development of portable automatic number plate recognition system on android mobile phone. IOP Conf. Ser. Mater. Sci. Eng. **53**, 012066 (2013)
18. Machaj, J., Brida, P., Benikovsky,J.: Scalability optimization of seamless positioning service. Mobile Inf. Syst. **2016** (2016)
19. Ajami, S.: Use of speech-to-text technology for documentation by healthcare providers. Natl. Med. J. India **29**(3), 148–152 (2016)
20. Kurtz, J.: ASP.NET MVC 4 and the web API: Building a REST service from start to finish, vol. 9781430249788 (2014)
21. Maresova, P., Sobeslav, V.: Effective evaluation of cloud computing investment—application of cost benefit method analysis. E. M. Ekon. Manag. **20**(2), 134–145 (2017)
22. Maresova, P., Sobeslav, V., Krejcar, O.: Cost–benefit analysis–evaluation model of cloud computing deployment for use in companies. Appl. Econ. **49**(6), 521–533 (2017)
23. Sobeslav, V., Balik, L., Hornig, O., Horalek, J., Krejcar, O.: Endpoint firewall for local security hardening in academic research environment. J. Intell. Fuzzy Syst. **32**(2), 1475–1484 (2017)
24. Zubr, P., Dvorak,J., Selamat, A., Krejcar, O.: Smart security system based on android platform. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 10486, pp. 115–128.LNCS (2017)
25. Sadiq, F.I., Selamat, A., Ibrahim, R., Selamat, M.H., Krejcar, O.: Stampede prediction based on individual activity recognition for context-aware framework sing sensor-fusion in a crowd scenarios. Front. Artif. Intell. Appl. **297**, 385–396 (2017)
26. Kozlovsky, J., Dvorak, J., Krejcar, O.: Location based services used in smart electronic wallet mobile application. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 9847, pp. 353–362. LNCS (2016)
27. Guerrero-Narváez, S., Niño-Zambrano, M., Riobamba-Calvache, D.-J., Ramírez-González, G.-A.: Test bed of semantic interaction of smart objects in the web of things. Futur. Intern. **10**(5) (2018)

28. Perneger, T.V., Courvoisier, D.S., Hudelson, P.M., Gayet-Ageron, A.: Sample size for pre-tests of questionnaires. Qual. Life Res. **24**(1), 147–151 (2015)
29. Laneuville, O., Sikora, D.: Quantitative analysis of the usage of a pedagogical tool combining questions listed as learning objectives and answers provided as online videos. Futur. Intern. **7**(2), 140–151 (2015)

Integration Challenges for Outsourcing of Logistics Processes in E-Commerce

Adam Wasilewski

Abstract Outsourcing of logistics processes is an interesting alternative to the traditional logistics model for e-shops. This approach can significantly reduce logistics costs and improve business efficiency. However outsourcing requires consideration of significant issues related to business operations and some organizational and technical problems. This study presents the differences in the approach to systems integration in the standard outsourcing model and in the dropshipping model, the advantages and disadvantages of both options as well as business and integration challenges for the implementation of the outsourcing in the e-commerce. Particular attention was paid to the scope of integration between IT systems, necessary for the proper functioning of any e-shop using outsourcing or dropshipping.

Keywords Logistics · E-commerce · System integration · Dropshipping

1 Introduction

Logistics in e-commerce is one of the most important factors influencing success or failure. The size of the offer, prices and delivery time depends on the chosen logistics model. On the other hand approach to logistic issues highly influence the costs of the business.

Typical logistics strategies in e-commerce include [2]:

- confectioning of orders and shipments from own warehouse,
- outsourcing the execution of orders to external logistics operators,
- dropshipping—outsourcing the execution of orders to suppliers (manufacturers, wholesalers).

The first approach (own warehouse) generates the biggest cost (including warehouse rental and employees) and is inflexible in the case of a change in the scale of

A. Wasilewski (✉)
Department of Computer Science and Management, Wroclaw University of Technology, Wroclaw, Poland
e-mail: Adam.Wasilewski@pwr.edu.pl

M. Huk et al. (eds.), *Intelligent Information and Database Systems: Recent Developments*, Studies in Computational Intelligence 830,
https://doi.org/10.1007/978-3-030-14132-5_29

operations. The main advantage of own warehouse is independence from external companies.

The second approach (outsourcing of logistics processes to specialized companies) transfers the responsibility for the storage of goods and the execution of orders to an external company. Usually payments for a logistics operator include a fixed part (resulting from storage fees) and a variable part (resulting from the number of products shipped). Such solution is economically balanced and flexible. Moreover, many logistics operators have big experience in e-commerce logistics and processes, including RMA (return merchandise authorization), and can help sellers to improve their business and customer service [4].

Dropshipping is an alternative for typical outsourcing of logistics processes and can be defined as “a form of logistic outsourcing in which the order fulfillment process is transferred from the online shop to its supplier (manufacturer, wholesaler)” [1]. This approach changes the responsibility for sending goods to the end customer. In previous strategies goods were packed and sent by online store or specialized company. Dropshipping transfers the responsibility for the fulfillment of an order (or part of it) to the manufacturer or wholesaler and can be used both in B2C (Business to Customer) and B2B (Business to Business) sales. Dropshipping let to minimize own stock (in extreme cases even to zero) but needs to define specific processes for post-sale service (including returns).

The aim of the study is to analyze different approaches to outsource logistics processes in e-commerce platform (which includes e-shop and external systems). Realization of the objective was based on literature, case studies and practical implementations based on business concepts of outsourcing implementations for online stores.

2 Similarities and Differences Between the Outsourcing Strategies

Both outsourcing strategies of logistics processes in e-commerce have a number of positive features compared in Table 1.

Outsourcing of logistics processes allow to:

- decrease warehousing costs because the online seller does not have to bear logistics costs related to goods purchase and storage,
- improve cashflow in dropshipping option—payment for the goods payment for goods is deferred until products are definitely sold (at least 14 days after sale because of law regulations),
- increase assortment in dropshipping option—possible contracts with large number of manufacturers and wholesalers allow for a significant increase in the availability of goods without incurring the costs of their purchase by the online store.

In addition to the positive features of outsourcing some negative features can be found (Table 2).

Table 1 Comparison of positive features of outsourcing and dropshipping

Feature	Outsourcing to specialized company	Dropshipping
Warehouse costs	Storage fees + fee per order (sent, returned etc.)	Order fee (can be multiplied if the customer's order has been split between multiple suppliers) + costs of internal or external post-service processes (returns etc.)
Stock	Limited by owned goods	Limited by goods in stock in suppliers' warehouses
Payments from customers	Final customers pay in advance or on delivery	Final customers pay in advance or on delivery
Payments to suppliers	Payment for the entire delivery of goods in accordance with the terms of purchase	Payment only for successfully sold products (after deduction of returns)
Flexibility	Usually long-term contract with logistics operator	Possible long-term and short-term contracts with many manufacturers and wholesalers

Table 2 Comparison potentially negative features of outsourcing and dropshipping

Feature	Outsourcing to specialized company	Dropshipping
Cost of delivery	Standard	May be multiplied by number of items in the order
Sale document	One sale document issued by logistics operator in the name of seller	One or many sale documents (per each package) issued by seller
Post-sale processes	Typical processes, carried out by logistics operator	Complex processes, carried out by seller
Shipment	One shipment	One or many shipments
Delivery time	Standard	Long

The main issues to be analyzed and solved before implementation of outsourcing strategy in e-commerce are:

- possible increased costs of transport to the customer in dropshipping option (if items from the order are delivered from different locations; in this case delivery costs may be higher than the cost of delivery incurred by the customer when placing an order in the online store),
- documentation of sales (one or many fiscal documents in the case of multiply deliveries? when to issue documents—after each shipment or after the entire order has been shipped?),

- handling of withdrawals and complaints (where customer should return products? who is responsible for customer service?),
- multiple shipments in dropshipping option (if the order is processed by multiple suppliers, the end customer will receive multiple shipments, which may be cumbersome and not accepted by all customers, e.g. due to the need to make multiple arrangements with couriers),
- delivery time (dropshipping option needs to split the order into suborders that takes time if systems integration is inefficient),
- cost of IT systems integration (effective use of dropshipping requires comprehensive integration of IT systems and integration must include all suppliers),
- reliability of stock levels (lack of own stock in dropshipping option makes the shop dependent on information about suppliers' stocks and reliable information on stock levels is key to ensuring high customer satisfaction).

One of the most important problem (especially in dropshipping option) is transparency of transactions. Splitting of the order into suborders may reduce transparency of transactions for the customer, especially if the split is not visible to the customer and the customer was not informed about it. Moreover goods from the one order may be delivered at different times, by different couriers and the sales document may be delivered regardless of the goods and in such case customer may be confused how to return goods if necessary.

3 Implementation of Typical Outsourcing in E-Commerce

There are two basic flows in online shops:

- the flow of goods offered in the shop,
- the flow of orders placed by final customers.

Other flows (e.g. financial or sale documents) are depend and closely linked to main flows.

Chosen logistics strategy significantly influence all process flows. The flows of orders and goods in the standard outsourcing model (order picking and dispatch from specialized company's warehouse) are shown in Fig. 1.

The order placed by the end customer in the online store is sent to the warehouse provided by specialized company (logistics operator). Such order is completed and sent (if the size allows it in one shipment) to the customer [3].

Availability of goods in the online store results from the actual stock in the warehouse. Stocks may change primarily in the case of order fulfilment (external issues) and deliveries from manufacturers or wholesalers (external receptions). Stocks are to a lesser extent influenced by internal warehouse operations (inter-warehouse shifts), which may be the result of product defects.

Typical integration architecture for standard outsourcing approach is shown in Fig. 2.

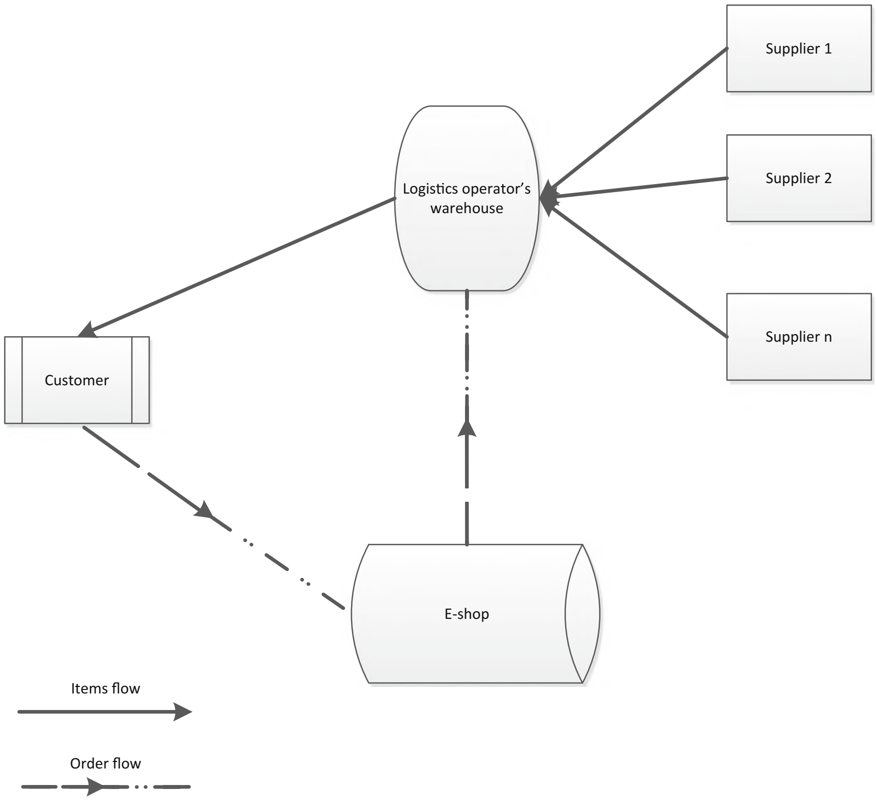


Fig. 1 Order and goods flow for standard outsourcing model. *Source* Own preparation

Systems' integration in the standard outsourcing model is typical for e-commerce solutions.

Online shop must be integrated with online payments and with WMS (Warehouse Management System). In addition WMS should send information about dispatches to Finance Management System (for fiscal purposes).

WMS is integrated with couriers to ensure the shipment of goods. The number of available shipping options determines the complexity of integration with couriers but usually it is limited number of delivery methods to choose from.

4 Implementation of Dropshipping in E-Commerce

Flows of orders and goods differ significantly when the dropshipping model is used. After placing an order, it is transferred to the supplier(s) and the supplier is responsible for packaging and sending the goods to the end customer. The online shop only issues

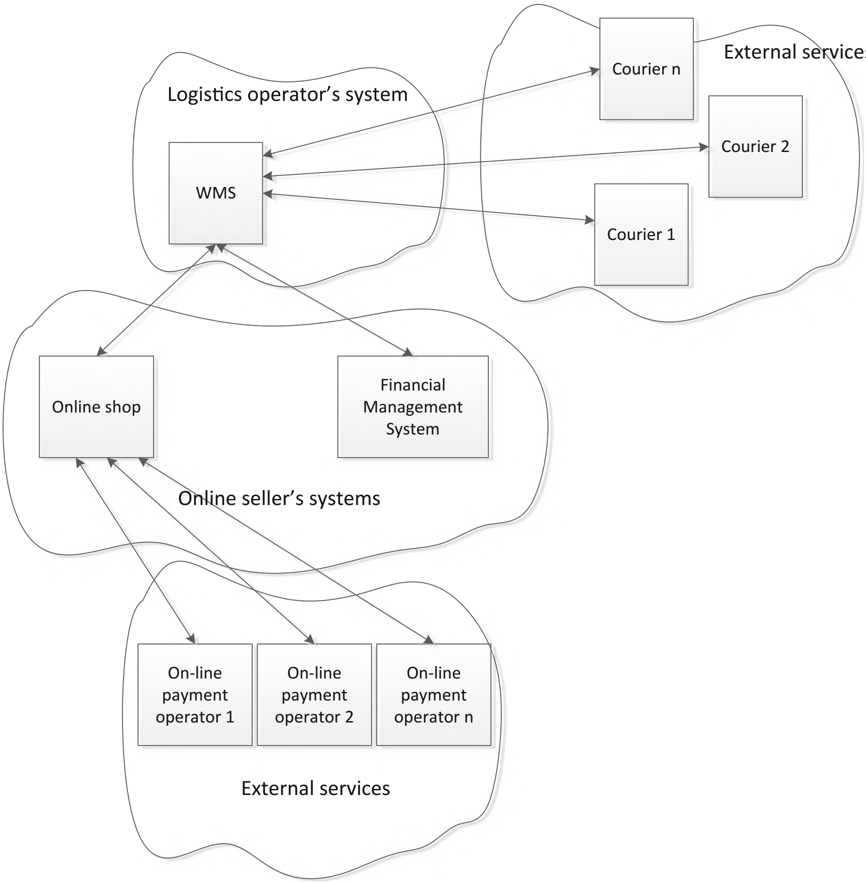


Fig. 2 Example of integration architecture for standard outsourcing approach. *Source* Own preparation

sale document (for fiscal purposes) and delivers this document to the end customer (it means that from the end customer point of view seller is the party to the transaction and bears its legal consequences), and then accounts with the supplier of the goods.

The flows of orders and goods in the online shop operating in the dropshipping model is shown in Fig. 3.

One of the most important problems in dropshipping approach appears when the order contains items from different suppliers. In such case the online shop must split the order and distribute suborders to the suppliers.

Suppliers, on the basis of the received suborders, complete the goods in their warehouses and send them to the end customer, excluding physical delivery of the goods to the online seller.

Each order should contain (regardless of the logistics model) the following information:

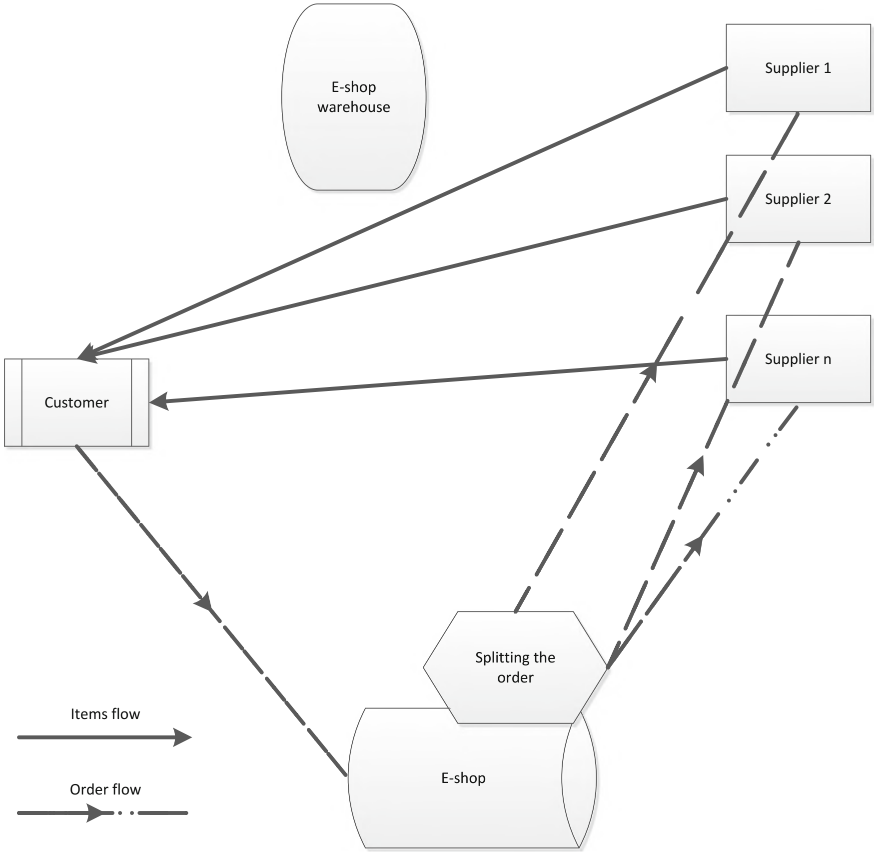


Fig. 3 Order and goods flow for dropshipping model. *Source* Own preparation

- header (basic information about the order, in particular its number, date of placing, delivery method, type of sale document),
- customer data (customer data may include information about the order, payer and recipient),
- items (ordered goods, quantities, prices, discounts granted),
- payment (method of payment, amount paid, date of payment, payment identifier),
- summary (delivery costs, total order value, comments).

For the purpose of suborders, it is sufficient to provide the suppliers with information about:

- address data of the recipient,
- the method of delivery,
- items to be shipped,
- the base order number.

Orders splitting requires an algorithm that optimize the fulfilment of the order in terms of:

- logistics costs (the cost of sending goods to the final customer should be as low as possible),
- delivery time (the delivery time of the entire order should be as short as possible),
- net margin (if the goods are available from multiple suppliers, the online seller's profit should be as high as possible),
- prioritization of suppliers (when an online seller wants to force an order to be sent to a particular supplier, e.g. to increase turnover with that supplier and obtain better terms and conditions).

In addition, the algorithm may take into account the preferences of the end customers—e.g. consent to a longer delivery time in exchange for a lower cost of delivery.

Effective implementation of dropshipping is not possible without the integration of IT systems. The minimum scope of integration must include:

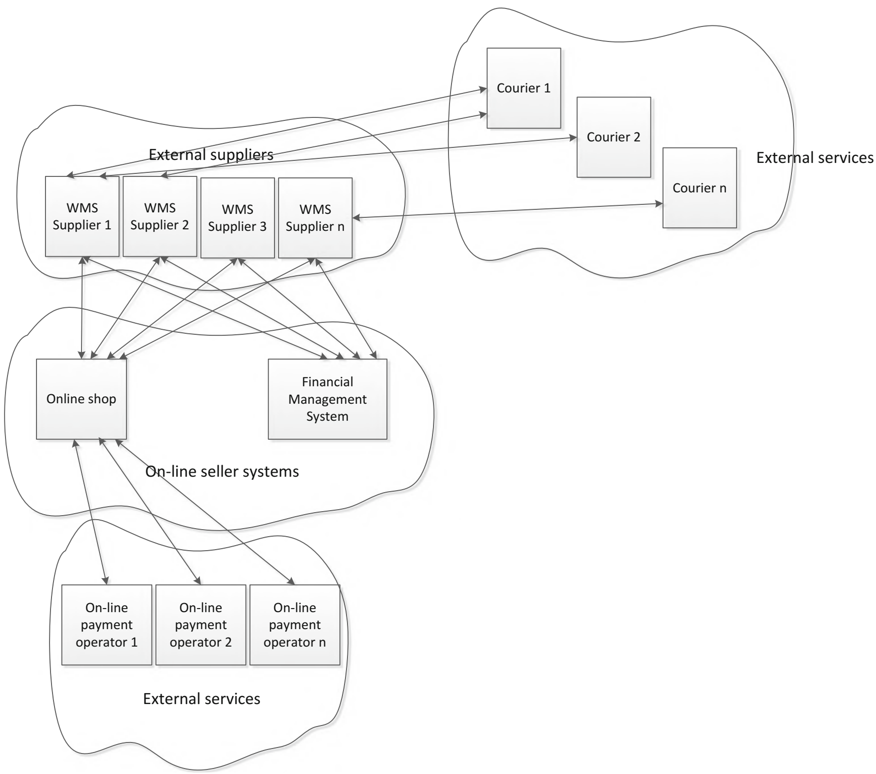


Fig. 4 Example of integration architecture for dropshipping approach with point-to-point integrations. *Source* Own preparation

- transfer of orders (and suborders),
- order statuses (including tracking number to track shipments),
- synchronization of stock levels (in real time if possible—to reduce the probability of placing an order for goods that are no longer in the supplier's possession).

Decision about an architecture of integration is crucial in dropshipping approach because of the number of systems that must be integrated. E-shop needs integration with each supplier separately and, theoretically, each supplier may have a different system, which affects both the functional and technical aspects of integration. In such circumstances standard point-to-point integration (shown in Fig. 4) may not be efficient.

Point-to-point integration in dropshipping approach seems to be extremely inflexible, expensive and hard to maintain. Such solution does not let to apply to main advantage of dropshipping—the possibility to quickly and easily start cooperation with new manufactures and wholesalers.

Effective implementation of dropshipping requires an open model of integration, allowing for quick and cheap integration of new suppliers. A possible solution is to use elements of the SOA (Service Oriented Architecture) paradigm, in particular ESB (Enterprise Service Bus) integration bus. This approach unifies the integration between the e-shop and ESB and enables the preparation of individual integrations as web services available through ESB (Fig. 5).

Systems' integration based on SOA let to reuse external services (e.g. couriers) and internal integration (e.g. with suppliers' WMS systems). If new supplier uses

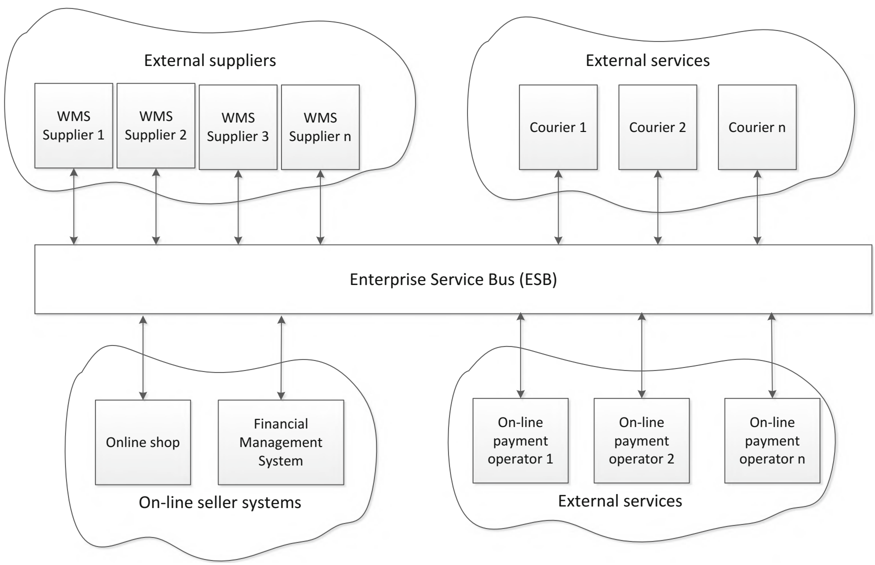


Fig. 5 Example of integration architecture for dropshipping approach with ESB. *Source* Own preparation

WMS system that was integrated before (because any previous suppliers has this system) then new integration can be very quick and not expensive.

5 Summary

Outsourcing is an interesting alternative to the traditional logistics model in online shops. Its main advantage is the possibility to reduce the costs of running a business. On the other hand it has some disadvantages and all the consequences must be consider rationally. The most important issues of the outsourcing of logistics processes in e-commerce were presented in this study, with particular emphasis on the requirements related to the integration of IT systems. Successful outsourcing requires the implementation of effective integration between IT systems to ensure high execution of orders placed by customers.

References

1. Chodak, G.: Wybrane zagadnienia logistyki w sklepach internetowych—modele, badania rynku. Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław (2014)
2. Kawa, A.: Fulfillment service in e-commerce logistics. LogForum **13**(4), 429–438 (2017)
3. Lang, G.: Multi-channel retail supply chain management: fulfillment systems in multi-channel retailing—customer expectations and economic performance, 2010. In: 8th International Research Conference in Logistics and Supply Chain Management (RIRL), Sep 2010, Bordeaux, France, pp. 1–25. Strategy and Supply Chain Management (2010)
4. Wasilewski, A.: Integracja systemów informatycznych w ramach farm e-commerce. In: Zeszyty Naukowe—Uniwersytet Szczeciński. Studia Informatica, nr 38, s. 215–226 (2015)

Performance Improvement of Open Source Based Business Intelligence System Using Database Modeling and Outlier Detection

Tsatsral Amarbayasgalan, Meijing Li, Oyun-Erdene Namsrai, Bilguun Jargalsaikhan and Keun Ho Ryu

Abstract With all the advanced technology nowadays, new data is being generated every minute. For example, the average size of the computer's hard disk is 10 gigabytes in 2000, today on the Facebook website has increased 500 terabytes of new data per day [1]. Data is growing rapidly, but it is not enough valuable. Thus, it is important to extract information that is useful in the future from a large amount of data. Business intelligence (BI) systems make a prediction that supports a business decision by analyzing collected data [2]. However, the accuracy of prediction depends on a data quality. In practice, data is usually a very low quality that includes many incomplete and anomaly data. Moreover, another problem is if data size increases, query response will be slow. Previous research work, we proposed a framework based on open-source technologies for the BI systems that possibility to analyze big data efficiently and apply it to the supermarket's BI system. Under this solution, we

T. Amarbayasgalan · B. Jargalsaikhan
Database Laboratory, School of Electrical and Computer Engineering,
Chungbuk National University, Cheongju 28644, Korea
e-mail: tsatsral@dblab.chungbuk.ac.kr

B. Jargalsaikhan
e-mail: bilguun@dblab.chungbuk.ac.kr

M. Li
College of Information Engineering, Shanghai Maritime University, Shanghai,
People's Republic of China
e-mail: mjli@shmtu.edu.cn

O.-E. Namsrai
School of Engineering and Applied Sciences, National University of Mongolia,
Ulaanbaatar, Mongolia
e-mail: oyunerdene@seas.num.edu.mn

K. H. Ryu (✉)
Faculty of Information Technology, Ton Duc Thang University, Ho Chi Minh City,
Vietnam700000
e-mail: khryu@tdtu.edu.vn; khryu@chungbuk.ac.kr

K. H. Ryu
Department of Computer Science, College of Electrical and Computer Engineering,
Chungbuk National University, Cheongju 28644, Korea

M. Huk et al. (eds.), *Intelligent Information and Database Systems: Recent Developments*, Studies in Computational Intelligence 830,
https://doi.org/10.1007/978-3-030-14132-5_30

have studied Hadoop data storage system, Hive data warehouse software, Sqoop data transmission tool and etc., successfully implemented them. In this paper, we have added anomaly detection stage on the proposed framework to improve information about related products that are purchased together by eliminating anomaly. Also, we have made an experimental study to improve the speed of time-dependent reports by applying the dimensional model to Hive data warehouse. In dimensional model data is stored in context of the single table (centralized context), and in relational model the context is distributed over many tables. As a result of the experimental study, the dimensional model is more efficient; its query response time is shown to be at least two times faster than the relational model based data warehouse.

Keywords BI system · Big data · Data warehouse · Data mining · Anomaly detection

1 Introduction

Since 1990, registration software was focused through the world. As a result of this, a sufficient amount of data was collected. Now, the amount of data is growing rapidly. Also, the size of raw data is big, but it is not enough valuable. Thus, instead of storing collected data as inactively, it is to predict the future by making a variety of processing on it. Data mining is the process of automatically discovering novel and useful patterns that might otherwise remain unknown in large data repositories [3]. In recent years, researches related to Decision support systems suggest using data mining for data analysis [4–7].

An abnormal pattern that is not compatible with most of the data in a dataset is named a novelty, outlier, or anomaly [8, 9]. The anomaly can be created for several reasons, such as data from different classes, natural variation, and data measurement or collection errors [3]. Data from a special situation, which is different from others, influences to generate the wrong results [10]. For example, we use data all citizens' income to estimate average income of "A" district. If the man who gets one billion paid, the average income of a particular district will be high. So after collection and problem definition, pre-processing is very important for data analysis [11]. For large and complex data, pre-processing is difficult and complicated that is a time-consuming task. It can be simplified using specialized data pre-processing tools.

Business organizations make a decision based on predicting that how to be useful in future and this prediction is generated by the BI systems. There are two main operations in the BI systems such as collecting and organizing a large amount of data from various sources and predicting by analyzing organized data. A data warehouse collects required data from a lot of data sources and thereby faces some problems that how to load data into the data warehouse and how to organize collected data.

We aimed to improve the architecture shown in Fig. 1. In other words, the previously research work proposed the approach, algorithm, and technologies required for store big data, create a data warehouse and extract knowledge from the data ware-

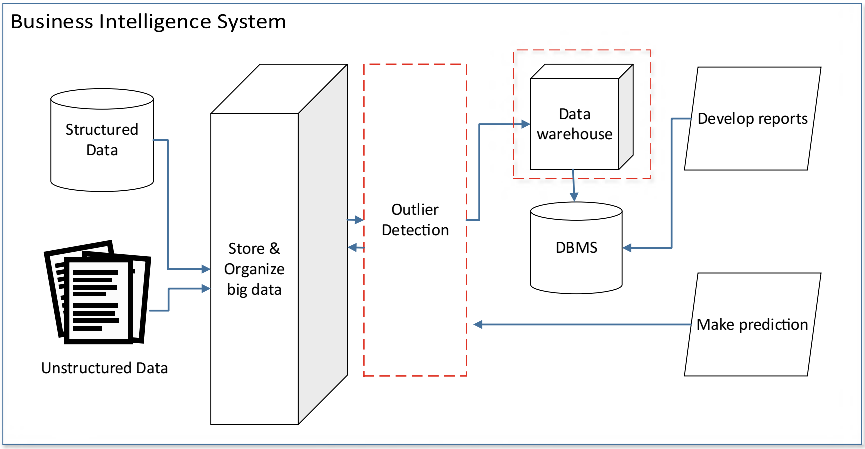


Fig. 1 General design of solving architecture

house in a comprehensive manner [12]. In this research work, we focused on how to improve the previous proposed framework by adding outlier detection layer and changing data warehouse organization schema.

2 Methodology

The architecture shown in Fig. 2 has been entirely designed to use Hadoop based open-source technologies. The advantage of Hadoop is that it can store both of structured and unstructured dataset. Also, Hadoop based open-source technologies are being invented and they are making more flexible usage of it. The proposed architecture in this research work has some advantages from the previous architecture by adding anomaly detection stage between the data storage and the prediction stages. Also, we suggested the dimensional model for the data warehouse to increase query processing time.

Firstly, data is collected to the HDFS file system on Hadoop from various data sources and then some required data for generating reports is loaded into the data warehouse. It is available to query the data from Hive data warehouse using HiveQL that is SQL like language. But taking summarized report from Hive data warehouse directly is slow (it takes a long time to develop each report due to processing big data). Thus, summarized results from Hive data warehouse are copied into MySQL database server, and then an end-user application retrieves reports from MySQL database server that spends less time. After collecting and storing dataset into Hadoop and Hive data warehouse, anomalies will be removed by outlier detection system before analyzing the collected dataset.

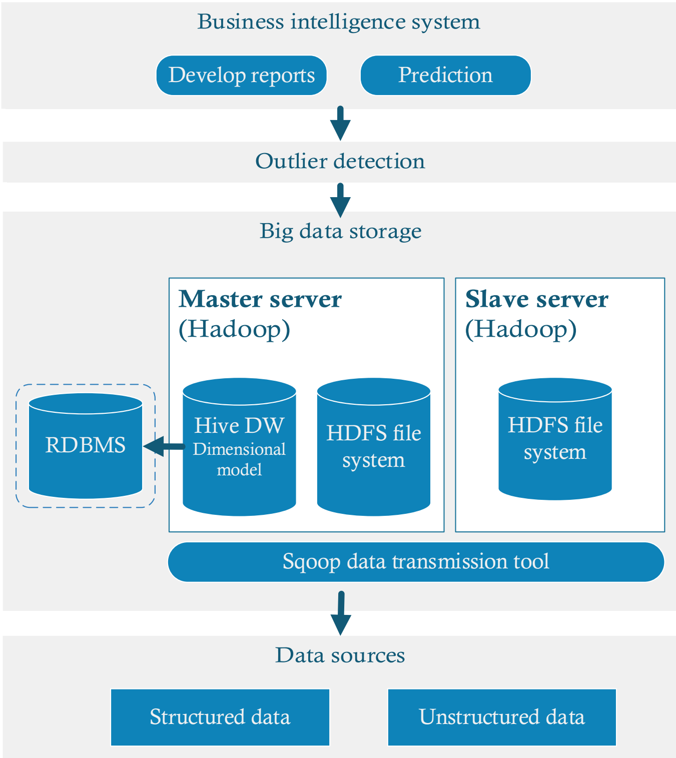


Fig. 2 The solution of methods and technologies of proposed architecture

2.1 *Big Data Storage*

We studied what technology is used to store big data for analysis.

2.1.1 Hadoop Data Storage Technology

Alfredo et al. [13], Rogers et al. [14] and Li et al. [15] are suggested Hadoop based big data storage for BI systems. Apache Hadoop is open-source software for reliable, scalable and distributed computing [16]. It is designed to scale up from a single server to thousands of machines, distribute large data sets across the cluster of machines. So, it detects and handles failures of machines and application that connected to a machine does not stop and continues to work. Hadoop has the following advantages:

- Scalable (It is available to add a new node into the cluster)
- Flexible (It can accept structured and unstructured data from various data sources)

- Reliable (System doesn't stop and continues to work when failures occur in any node)
- Parallel processing (It can work on simple performance computers)
- Open-source.

2.1.2 Hive Data Warehouse Software

Hive open-source data warehouse software is based on Hadoop. It facilitates querying and managing large datasets residing in distributed storage [17]. Hive data warehouse store both of structured and unstructured data and query the data using a SQL like a language called HiveQL.

Hive data warehouse is accessed with three methods including CLI (Command line interface), web interface and thrift server (using JDBC or ODBC). CLI and web interface opportunities for direct access to a warehouse that is not passed by a user application. But thrift server must be started for access to the data warehouse on Hive server from the user application. This server provides to get information from Hive data warehouse by sending HiveQL and using any programming language from a remote user. Under the research work, we have loaded the supermarket's basket data (transaction data) into Hadoop system, then created Hive data warehouse from it and received reports from the data warehouse.

2.1.3 Sqoop Data Transmission Tool

Sqoop successfully graduated from the Apache in March of 2012 and is now a top level Apache project. It is a tool designed for efficiently transferring bulk data between Apache Hadoop and structured data stores such as relational databases [18]. It works in two directions that are from and to database management system as well as from and to Hadoop. In other words, the Sqoop load data in the relational database is accessed from and to HDFS Hive and HBase.

2.2 Comparison Between Database Relational and Dimensional Models

Database dimensional model uses the concept of the relational model, but its connection schema between tables is different. Tables can connect with many tables in the relational model, but tables connect to only one center table in the dimensional model. In other words, all data is stored in context of the single table in dimensional model. However, context is dispersed over many tables in relational model. Thus, dimensional model does a few joining operations when extracting useful information

Table 1 Comparison between relational and dimensional model

Relational model	Dimensional model
It is a very complex model that make connections between the hundreds of tables	It is a very simple model that dimension tables connect to only one fact table
Flexible	Inflexible
Easy to update data effectively	Easy to get data effectively, less connected to each other
Data is temporary (it is modified several times)	Data is not temporary
Detailed data of operations	Aggregation result of batch operation for business decision
Simple reports	User friendly and various reports

from it and executes fast. Data can be stored effectively by the relational model and data can be extracted effectively by dimensional model.

Dimensional model consists of two kinds of table such as fact and dimension [13]. The central table is the only table in the schema with multiple joins connecting it to all the other tables. This central table is called the fact table and the other tables are called dimension tables. Numeric values of business operations are stored in fact table and text values are stored in dimension tables [14]. Time dimension table is used to query based on time. Table 1 is shown a difference between relational and dimensional models:

There are following advantages by organizing Hive data warehouse by dimensional model:

- Simple query—joining operation in the dimensional model is simpler than joining operation in the relational model
- Simplified business report—generate a time-dependent report from the dimensional model is easier than the relational model
- Query performance—data can be extracted effectively by the dimensional model. Because tables in the dimensional model have a fewer relationships and a query makes a less join operations.

2.3 Anomaly Detection

Belongs to the supermarket's BI system, a binary dataset is generated from a transactional dataset for identifying frequently together purchased products. We used mean and standard deviation to detect anomaly from the binary dataset which is transaction dataset. Algorithm 1 shows the anomaly detection from binary dataset step by step:

Algorithm 1: Anomaly detection

Input: List of products X, $\{X_1^k\}$, List of transactions T, $\{T_1^n\}$
months ← all months in T
for i **to** X **do**
    std_i ← Standard deviation of the product i
    avg_i ← Average value of the product i
    $count_i$ ← []
    for m **to** months **do**:
        for t **to** T **do**:
            if i **in** t:
                $count_i$[month] ++
        if $(std_i \pm avg_i)$ < count **or** count < $(std_i \pm avg_i)$:
            Remove i from all transactions during m
end for

3 Experimental Study

In this section, we have made two kinds of an experiment such as comparison of query response time between the data warehouses that are organized by dimensional and relational models, and anomaly detection. Two kinds of datasets were used in our experiment such as a real dataset with 200,057 records for generating reports and a synthetic dataset with 100 records for outlier detection.

3.1 Comparison Between Dimensional and Relational Model Based Data Warehouses

In this paper, we improved reporting time by replacing the relational model of the Hive data warehouse with the dimensional model. On the other hand, we have performed the experiment by creating two separate data warehouses to store the same dataset by two different modeling that are relational and dimensional. Then we tested which model is less query response time. Figure 3 shows a difference between the organization of relational and dimensional models for the supermarket's data warehouse that stores information about supermarket's branch, POS, cashier, product category, product, purchase, and bill.

Figure 4 shows the folders that store data of dim_branch, dim_cashier, dim_pos, dim_product, dim_time, dim_category and fact_sales tables on the HDFS file system of Hadoop. Also, we created folders for data warehouse which is organized by the relational model.

After creating all required data folders on HDFS file system of Hadoop, some folders that will be used in data warehouse were connected to the corresponding tables in Hive data warehouses. By connecting folders on Hadoop into Hive data warehouse, future reports can be developed by HiveQL query language. Figure 5 shows tables

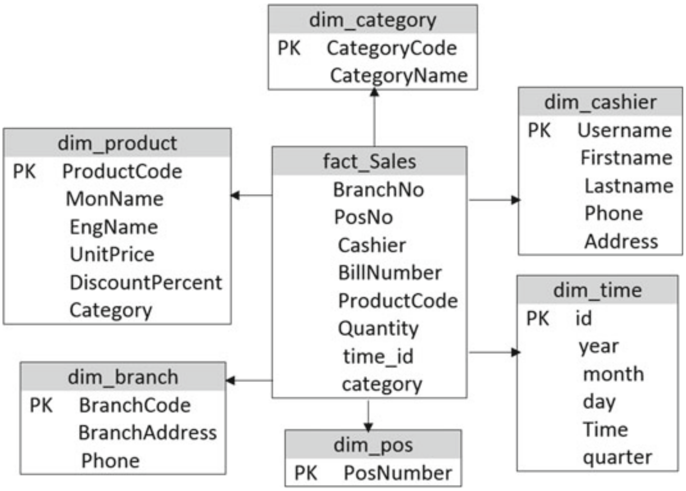
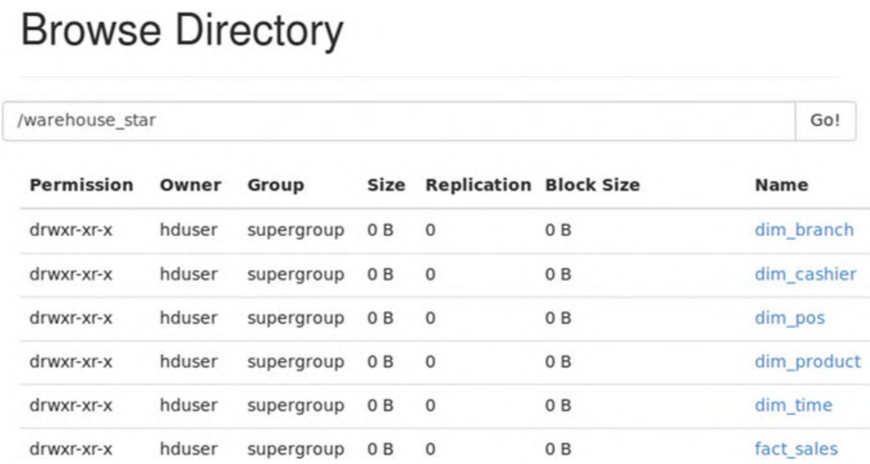


Fig. 3 Dimensional model of Hive data warehouse



Browse Directory

/warehouse_star Go!

Permission	Owner	Group	Size	Replication	Block Size	Name
drwxr-xr-x	hduser	supergroup	0 B	0	0 B	dim_branch
drwxr-xr-x	hduser	supergroup	0 B	0	0 B	dim_cashier
drwxr-xr-x	hduser	supergroup	0 B	0	0 B	dim_pos
drwxr-xr-x	hduser	supergroup	0 B	0	0 B	dim_product
drwxr-xr-x	hduser	supergroup	0 B	0	0 B	dim_time
drwxr-xr-x	hduser	supergroup	0 B	0	0 B	fact_sales

Fig. 4 Folders in the HDFS file system that store data of tables in the data warehouse organized by dimensional model

in the data warehouse organized by dimensional model. We used HiveQL to show list of tables in the data warehouse on Hive by CLI.

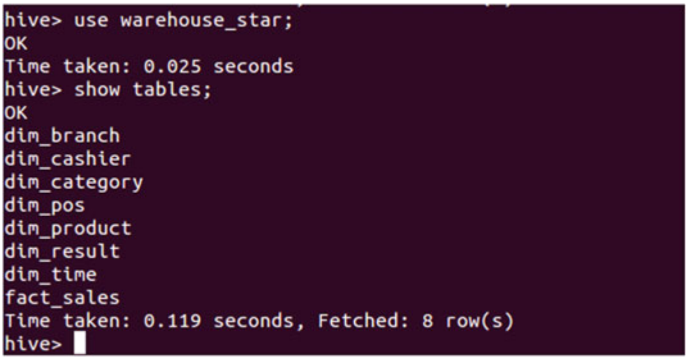


Fig. 5 Tables in data warehouse organized by dimensional model

3.2 Anomaly Detection

We used the supermarket's synthetic dataset for testing our anomaly detection system before discovering related products that are purchased together. This dataset has 8 attributes (date, apple, coca cola, toothbrush, carrot, potato, kids book, onion) and 100 instances. We have discovered related products from two kinds of a dataset which are anomaly removed and anomaly not removed using the FP-Growth algorithm that implemented previous research work. Then these two results have been compared.

4 Result

In this section, we present the result of comparison between data warehouses and implemented anomaly detection system by proposed framework solution.

4.1 Result of Query Processing Time on Data Warehouses

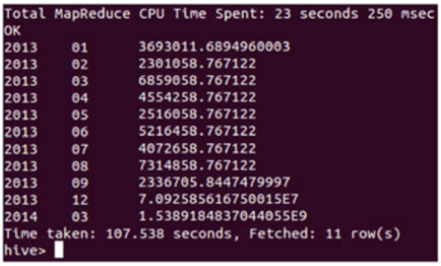
To compare the results of data warehouses that organized by dimensional and relational models, we have created two types of data warehouse in the Hive and copied the same dataset into these warehouses. After creating required folders on HDFS file system in Hadoop, tables that store XXX supermarket's purchase information from 2013 to 2014 in MySQL server have been imported into these created folders using Sqoop data transmission tool. As a result of this step, 200,057 purchase records are copied into 2 different Hive data warehouses.

We run queries that return the same result on these data warehouses and compare speed and query syntax. We run a query that takes an income of each month from

```
select r.year, r.month, sum(r.total) from (select sub-
string(b.date, 1, 4) as year, substring(b.date, 5, 2) as
month, b.total from (select a.date, a.time, a.number,
sum(a.quantity*p.unitprice) as total from purchase as a,
product as p where a.productcode = p.productcode group by
date, time, number) as b) as r group by r.year, r.month;
```

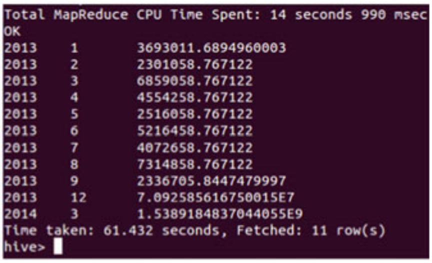
Fig. 6 The select query that retrieves an income of each month from data warehouse organized by the relational model

(a)



```
Total MapReduce CPU Time Spent: 23 seconds 250 msec
OK
2013    01    3693011.6894960003
2013    02    2301058.767122
2013    03    6859058.767122
2013    04    4554258.767122
2013    05    2516058.767122
2013    06    5216458.767122
2013    07    4072658.767122
2013    08    7314858.767122
2013    09    2336705.8447479997
2013    12    7.092585616750015E7
2014    03    1.5389184837044055E9
Time taken: 107.538 seconds, Fetched: 11 row(s)
hive>
```

(b)



```
Total MapReduce CPU Time Spent: 14 seconds 990 msec
OK
2013    1     3693011.6894960003
2013    2     2301058.767122
2013    3     6859058.767122
2013    4     4554258.767122
2013    5     2516058.767122
2013    6     5216458.767122
2013    7     4072658.767122
2013    8     7314858.767122
2013    9     2336705.8447479997
2013    12    7.092585616750015E7
2014    3     1.5389184837044055E9
Time taken: 61.432 seconds, Fetched: 11 row(s)
hive>
```

Fig. 7 The result of the select query that retrieve the monthly report from data warehouse. **a** data warehouse is organized by the relational model, **b** data warehouse is organized by the dimensional model

```
select B.year, B.month, sum (B.total) from (select t.year,
t.month, a.* from (select f.*, p.unitprice*f.quantity as to-
tal from fact_sales as f, dim_product as p where
f.productcode = p.productcode) as a, dim_time t where
a.time_id = t.id) as B group by B.year, B.month;
```

Fig. 8 A select query that retrieves an income of each month from data warehouse organized by the dimensional model

data warehouse which organized by relational model, and this query syntax is shown in Fig. 6. As a result, the total income of each month from January 2013 to March 2014 was retrieved and spent 107.538 s for query processing. Figure 7a shows the result that includes income of 11 months from the Hive data warehouse organized by the relational model using the CLI.

After that, we run a query that retrieves the same result with the query shown in Fig. 7a from the data warehouse organized by the dimensional model, and query syntax is shown in Fig. 8. As a result, shown in Fig. 7b, the total income of each month from January 2013 to March 2014 was retrieved and spent 61.432 s for query processing that shown in Fig. 8.

We run 3 kinds of query that take the total income of each month, quarter, and year. By month, the query of the data warehouse organized by the relational model

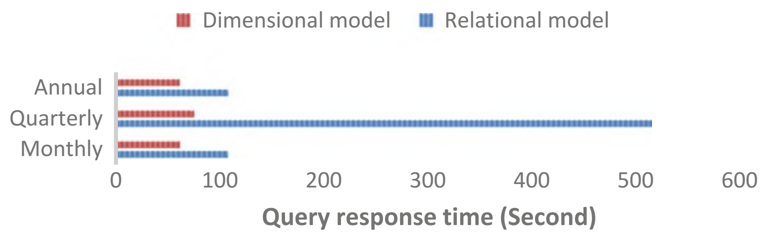


Fig. 9 Comparison of query processing time for time-dependent reports between the dimensional and relational models based warehouses

spent 107.538 s and the query of the data warehouse organized by the dimensional model spent 61.432 s. By quarter, the query of the data warehouse organized by the relational model spent 515.16 s and the query of the data warehouse organized by the dimensional model spent 75.131 s. By annual, the query of the data warehouse organized by the relational model spent 61 s and the query of the data warehouse organized by the dimensional model spent 107.84 s. As a result of the test, the query of the dimensional model could work simpler and faster than the relational model. Figure 9 shows the comparison of query processing time between 2 differently organized data warehouses shown in Fig. 3. The centralized context based organization is faster than distributed organization when to take a summarized result from data warehouse.

4.2 *Result of Anomaly Detection Test*

In this section, we present a result of implemented anomaly detection system by the proposed solution. Figure 10 shows the graphical user interface of the anomaly detection system. The red box refers to the purchase of apple. Apple was purchased constantly each month except May, but it increased rapidly in May due to Children's Day. Therefore, purchased records of apple in May is the anomaly.

Figure 11 shows 2 different results of related products from datasets with and without anomaly. The confidence and minimum support were configured as 96% and 0.01, respectively. According to the result shown on the left side of Fig. 11, kid's book and apples are sold together frequently. But this rule is valid only Children's Day and not valid daily. The right sight of Fig. 11 shows related products from dataset without anomaly. The confidence and minimum support were configured the same with previous processing. According to result, kid's book and apple are not related to each other. The discovered result has been improved by remove anomaly.

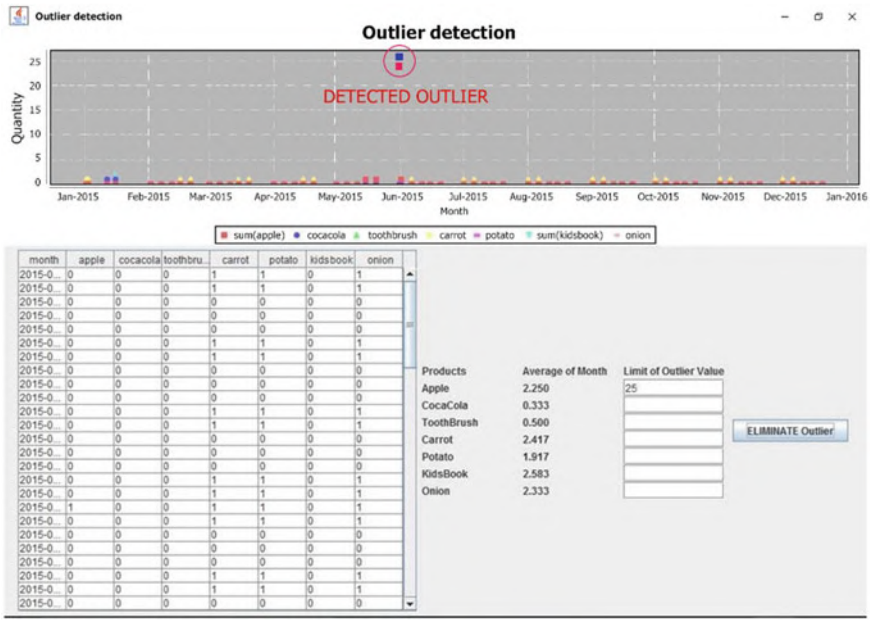


Fig. 10 Anomaly detection system

```
Best rules found:

 1. onion=T 28 ==> carrot=T 28    <conf:(1)> lift:(3.1)
 2. potato=T 23 ==> carrot=T 23    <conf:(1)> lift:(3.1)
 3. potato=T 23 ==> onion=T 23    <conf:(1)> lift:(3.21)
 4. potato=T onion=T 23 ==> carrot=T 23    <conf:(1)> li
 5. carrot=T potato=T 23 ==> onion=T 23    <conf:(1)> li
 6. potato=T 23 ==> carrot=T onion=T 23    <conf:(1)> li
 7. carrot=T 29 ==> onion=T 28    <conf:(0.97)> lift:(3.
 8. apple=T 27 ==> kidsbook=T 26    <conf:(0.96)> lift:(
```

```
Best rules found:

 1. onion=T 28 ==> carrot=T 28    <conf:(1)> lift:(3.1
 2. potato=T 23 ==> carrot=T 23    <conf:(1)> lift:(3.
 3. potato=T 23 ==> onion=T 23    <conf:(1)> lift:(3.2
 4. potato=T onion=T 23 ==> carrot=T 23    <conf:(1)>
 5. carrot=T potato=T 23 ==> onion=T 23    <conf:(1)>
 6. potato=T 23 ==> carrot=T onion=T 23    <conf:(1)>
 7. carrot=T 29 ==> onion=T 28    <conf:(0.97)> lift:(
```

Fig. 11 Results of related products from datasets with and without anomaly. Left side shows the result from dataset with anomaly and the right side shows result from dataset without anomaly

5 Conclusion

In previous research work, we have proposed the solution of how to implement the BI system to generate decision support information efficiently from big data and apply it to the supermarket's BI system. One of the advantages of our implementation was cost saving because we have used open-source technologies our configuration is simpler than commercial framework and tools; we aimed to find minimal time and space complexity solution, but its installation and configuration is more difficult than commercial software. Also, it has a lot of low-level configurations. In this research work, we have improved architecture of the previous research work by organizing data warehouse by dimensional model and adding anomaly detection step. We used simple data sets to test system components, and made 2 kinds of the experiment such as compare query response time and anomaly detection. In the comparison of query response time, we used a real dataset with 200,057 records but in outlier detection system, we used a synthetic dataset with 100 records. As a result of the experiment, a query of the dimensional model could work simpler and faster than the relational model. Also, we showed anomaly detection system can improve the true result of related products by experiment.

Acknowledgements This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (No. 2017R1A2B4010826), by the Private Intelligence Information Service Expansion (No. C0511-18-1001) funded by the NIPA (National IT Industry Promotion Agency), by the Business for Cooperative R&D between Industry, Academy, and Research Institute funded Korea Small and Medium Business Administration in 2017 (Grants No. C0541451) and also by the National Natural Science Foundation of China (61702324) in People's Republic of China.

References

1. Big Data Explained: [Online]. Available: https://www.mongodb.com/BIG-DATA-EXPLAINED
2. Chee, T., Chan, L.K., Chuah, M.H., Tan, C.S., Wong, S.F., Yeoh, W.: Business intelligence systems: state-of-the-art review and contemporary applications: In: Symposium on Progress in Information & Communication Technology, pp. 16–30, Kuala Lumpur, Malaysia (2009)
3. Tan, P.N., Steinbach, M.: Introduction to Data Mining, 1st edn. Pearson Education Inc., Boston (2006)
4. Shaw, M.J., Subramaniam, C., Tan, G.W., Welge, M.E.: Knowledge management and data mining for marketing. Decis. Support Syst. **31**(1), 127–137 (2001)
5. Shim, J.P., Warkentin, M., Courtney, J.F., Power, D.J., Sharda, R., Carlsson, C.: Past, present, and future of decision support technology. Decis. Support Syst. **33**(2), 111–126 (2002)
6. Lee, D.G., Ryu, K.S., Bashir, M., Bae, J.W., Ryu, K.H.: Discovering medical knowledge using association rule mining in young adults with acute myocardial infarction. J. Med. Syst. **37**(2), 9896 (2013)
7. Cho, Y.S., Kim, Y.A., Moon, S.C., Park, S.H., Ryu, K.H.: Effective purchase pattern mining with weight based on FRAT analysis for recommender in e-commerce. Lecture Notes in Electrical Engineering **330**, 443–454 (2015)
8. Chandola, V., Banerjee, A., Kumar, V.: Anomaly detection: a survey. ACM Comput. Surv. **41**(3), 1–58 (2009)
9. Amarbayasgalan, T., Jargalsaikhan, B., Ryu, K.H.: Unsupervised novelty detection using deep autoencoders with density based clustering. Appl. Sci. **8**(9), 1468 (2018)
10. Wu, X., Zhu, X., Wu, G.Q., Ding, W.: Data mining with big data. IEEE Trans. Knowl. Data Eng. **26**(1), 97–107 (2014)
11. Liu, K., Zhou, X., Feng, Y., Liu, J.: Clinical data preprocessing and case studies of POMDP for TCM treatment knowledge discovery. In: 14th International Conference on e-Health Networking, Applications and Services (Healthcom), pp. 10–14, IEEE (2012)
12. Amarbayasgalan, T., Bukhsuren, E., Namsrai, O., Ryu, K.H.: the approach of implementing business intelligence system: possibility to analyze big data. JARDCS (2), 775–779 (2018)
13. Alfredo, C.: Analytics over big data: exploring the convergence of data warehousing, OLAP and data-intensive cloud infrastructures. In: IEEE 37th Annual Computer Software and Applications Conference (2013)

14. Rogers, S.: Big data is scaling BI and analytics. Inf. Manage. **21**(5), 14–20 (2011)
15. Li, D., Park, H.W., Batbaatar, E., Munkhdalai, L., Musa, I., Li, M., Ryu, K.H.: Application of a mobile chronic disease health-care system for hypertension based on big data platforms. J. Sens. (2018)
16. Apache Hadoop https://hadoop.apache.org
17. Apache Hive https://hive.apache.org
18. Apache Sqoop http://sqoop.apache.org/

Designing an Intelligent Controller for Improving PEM Fuel Cell Efficiency

S. Dhanya, Rani Thottungal, Varghese Paul and S. Anand Hareendran

Abstract Maintaining the optimum performance of a PEMFC over a wide range of operating conditions is one of the greatest challenges in developing efficient and high performing fuel cell systems. This paper presents the effectiveness of neural network based intelligent controllers in regulating the partial pressure of Hydrogen, Oxygen and water under dynamic load conditions. Optimum values for these parameters are obtained using machine learning techniques on the simulation data obtained from the mathematical model of a PEMFC under various operating conditions of pressure, flow rate and humidity at the anode and cathode side of the PEMFC. Windrow Hoff algorithm and Neuro-Fuzzy controllers are used to attain the desired electrical performance of 25 V at 20 A from a 500 W PEMFC system. Genetic algorithm is basically used for validating the performance of the system. Thus an intelligent controller for optimum performance can be designed. The experimental results validates the process.

Keywords Electrolyte membrane fuel cell · Windrow Hoff algorithm · Neuro-Fuzzy controller · Intelligent controller

Supported by Muthoot Institute of Technology and Science, Kochi.

S. Dhanya
Department of ECE, Muthoot Institute of Technology and Science, Kochi, India
e-mail: dhanyas@mgits.ac.in

R. Thottungal
Department of ECE, Kumaraguru College of Technology, Coimbatore, India
e-mail: ranithottungal@yahoo.com

V. Paul
Department of CSE, Toc H Institute of Science and Technology, Kochi, India
e-mail: vp.itcusat@gmail.com

S. Anand Hareendran (✉)
Department of CSE, Muthoot Institute of Technology and Science, Kochi, India
e-mail: anandhareendrans@mgits.ac.in

M. Huk et al. (eds.), *Intelligent Information and Database Systems: Recent Developments*, Studies in Computational Intelligence 830,
https://doi.org/10.1007/978-3-030-14132-5_31

1 Introduction

Fuel cells are emerging as a promising source of green energy for the future over a wide variety of applications. The ability of PEMFCs to operate in under low temperature conditions, low corrosion, low weight and quick start up enables them to be deployed for various applications like automobile [1, 2], stationery power back up [4], portable power generation [3] compared to other fuel cell types like Solid Oxide Fuel Cells (SOFC), Molten Carbonate Fuel Cell (MCFC), Phosphoric Acid Fuel Cell (PAFC) and Alkaline Fuel Cells (AFC). Efficiency of controllers deployed on various PEMFC subsystem actuators is a major factor affecting the overall fuel cell system efficiency. Traditional control systems are good in controlling dynamic systems which can be represented using difference and differential equations. But for systems like PEM fuel cells which are highly nonlinear and dynamic in nature with large number of closely and dynamically coupled parameters, these traditional controllers become ineffective under various operating conditions. Hence there is a need for developing highly efficient intelligent controllers with adaptive nature which can handle nonlinear behaviour of PEMFCs effectively and thus improving the performance and efficiency of PEMFC. These intelligent controllers can cope with significant level of uncertainty as well as imprecision in the PEM fuel cell model and input parameter set and is capable of providing an optimum set of operating conditions which can give the desired electrical performance.

Intelligent control systems are developed by emulating the biological intelligence from nature with strong mathematical background and are highly efficient in complex problem solving.The fundamental methodologies used in implementing intelligence to control systems are fuzzy logic, neural networks and genetic algorithms. Procedures and algorithms developed using these techniques proves to be highly efficient in handling complex systems like PEMFC which is difficult to be represented using conventional control language.

This paper presents the design of Artificial Neural Network based intelligent controllers to regulate the values of most significant set of input parameters. The paper is organised as the following, Sect. 2 explains the importance of simulation study in understanding the dynamics of PEMFC and the advantage of using intelligent controllers to achieve desired electrical characteristics. Section 3 details the need for validating the range of input parameters which provide the optimum performance. Section 4 describes the implementation details of Windrow Hoff and Neuro-Fuzzy controllers. Results and discussions are included in Sect. 5.

2 Simulation Study and the Need for Intelligent Controller

The mathematical model of PEMFC in Matlab/Simulink is explained in this section. The reversible open circuit voltage (Nernst voltage) from a hydrogen Fuel cell is given as:

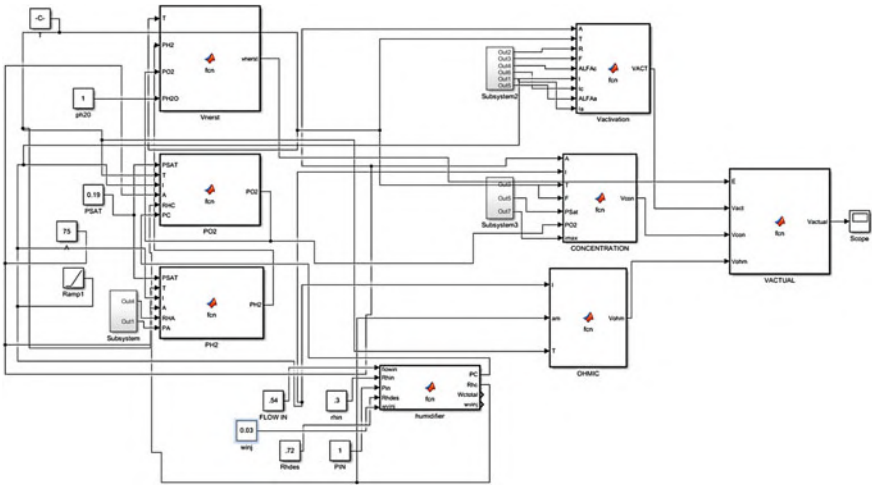


Fig. 1 Layout of matlab model for 500 W PEMFC

$$V_{reverse} = 1.229 - 0.85 * 10^{-3}(T_{fc} - 298.15) + 4.3085 * 10^{-5} T_{fc}[ln(p_{H2} + 0.5ln(p_{O2})] \quad (1)$$

The fuel cell output voltages is affected by various losses, Activation loss V_{act} is the voltage lost in driving the chemical reaction at the electrodes given by:

$$V_{act} = \left[\frac{RT}{4F\alpha c}\right] * ln\left(\frac{i}{ic}\right) * \left[\frac{RT}{2F\alpha a}\right] * ln\left(\frac{i}{i\alpha}\right) \quad (2)$$

Due to the flow of electrons through the electrodes and various interconnections, an electrical resistance is formed which includes the hindrance in the flow of ions in the membranes and results in a voltage drop proportional called Ohmic Loss. Decrease in stack voltage due to reduction in concentration of reactants at the electrode surface is known as mass transport loss (Concentration loss).

A mathematical model in Matlab/Simulink is developed by considering these dynamics of operation for a rated power of 500 W with 25 V at 20 A whose layout is shown in Fig. 1. Simulations are carried out to study the optimum range of input parameters for obtaining the optimum electrical response at the rated power.

The PEMFC model data and electrical performance under given set of operating conditions was validated by comparing with the specification manual of commercially available Horizon 500 W and the comparison plot is given in Fig. 2.

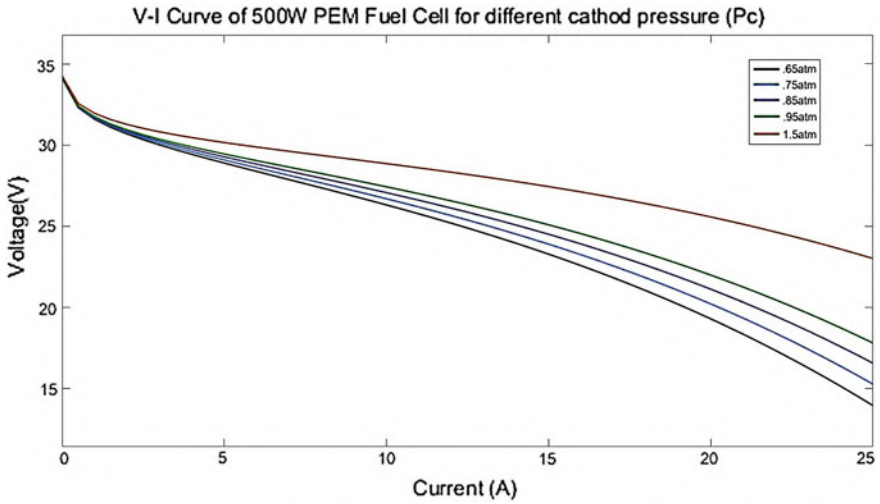


Fig. 2 Effect of Cathode Pressure On Stack Voltage

3 Implementation Details of Windrow Hoff and Neuro-Fuzzy Controllers

Results of simulation under a wide range of input conditions describe the effect of these parameters on stack performance and the suitable range of values to achieve the rated power. There is a highly nonlinear and inter related coupling between these input parameters which affects the stack voltage and hence the power at different load conditions.

The desired values of cathode inlet pressure, flow and relative humidity are obtained using three Neuro fuzzy Controllers:

1. To regulate $P_{a,co}$ using load current as input parameter at the cathode compressor valve.
2. To regulate W_{co} using load current as input parameter at the motor.
3. To regulate W_{inj} using $P_{a,co}$ and W_{co} as input parameters at the humidifier valve.

Windrow Hoff Algorithm

Least mean squares (LMS) algorithms are a class of adaptive filter used to mimic a desired filter by finding the filter coefficients that relate to producing the least mean squares of the error signal (difference between the desired and the actual signal). It is a stochastic gradient descent method in that the filter is only adapted based on the error at the current time. Windrow Hoff Algorithm is a family of LMS algorithm (Fig. 3).

Algorithm 1: Widrow-Hoff

initialize $\boldsymbol{w}_1 = \boldsymbol{0}$;
for $t = 1$ **to** T **do**
  get $\boldsymbol{x}_t \in \mathbb{R}^n$;
  predict $\hat{y}_t = \boldsymbol{w}_t \cdot \boldsymbol{x}_t$;
  observe y_t;
  incur loss of $(\hat{y}_t - y_t)^2$;
  update $\boldsymbol{w}_{t+1} = \boldsymbol{w}_t - \eta(\boldsymbol{w}_t \cdot \boldsymbol{x}_t - y_t)\boldsymbol{x}_t$;
end

Fig. 3 Windrow Hoff algorithm

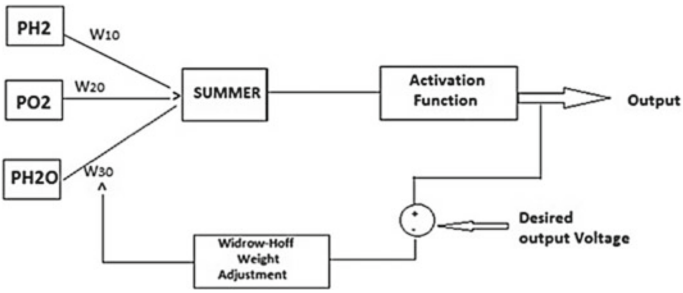


Fig. 4 Weight adjustment using Windrow Hoff algorithm

Give initial weight W_{i0} as 0 and n = 0. Now the input partial pressure P(n) is given by (Fig. 4),

$$P_n = W_{in} * P_i(n) \tag{3}$$

The new weight w_{in+1} is given by,

$$W_{in+1} = W_i(n) + \mu[Vd(n) - P_i(n)] * P_i(n) \tag{4}$$

Neuro Fuzzy controller

For the logic systems using fuzzy, which can be used with imprecise information, are good for explaining their decisions but they cannot automatically acquire the rules they use to make those decisions. These limitations have been a central driving force behind the creation of intelligent hybrid systems where two or more techniques are combined in a manner that overcomes the limitations of individual techniques. In order to overcome the problems associated with of knowledge acquisition, the neural networks can be used to automatically extract fuzzy rules from numerical data. For neural networks, the knowledge is automatically acquired by the back propagation algorithm, but the learning process is relatively slow and analysis of the trained network is difficult (black box). Neither is it possible to extract structural knowledge (rules) from the trained neural network, nor can we integrate special information about the problem into the neural network in order to simplify the learning procedure. In order to make a system which deals with cognitive uncertainties in a manner more like humans, one may incorporate the concept of fuzzy logic into the neural networks.

The computational process envisioned for fuzzy neural systems is as follows. It begins with the development of a fuzzy neuron based on the biological neuronal morphologies, followed by learning mechanisms.

A controller is required to maintain availability of oxygen to avoid oxygen starvation and thus extend the life of the fuel cell. Oxygen starvation is a critical condition that occurs when the partial pressure of oxygen falls below a certain value of the air stream in the cathode. This phenomenon results a rapid decrease in stack voltage, which in some cases can cause a hot spot, or even can cause burn-through of membrane. Although the oxygen starvation is spatially varying, this situation can be avoided by controlling the pressure and flow rate at the cathode.

Rule Steps

- Load data for training, testing, and checking.
- Generate an initial FIS model
- Choose the FIS model parameter optimization method
- Enter the number of training epochs
- Enter the training error tolerance
- View the FIS model output versus the training, checking, or testing data output

Identification of optimal parameters for improving efficiency always involve a huge cost and search component. Here a neuro learning system, derived from the genetic algorithm is used to explore the possibilities in improving the same. PEM fuel cell is the primary subject under the investigation. An optimal combination of A,B and C (that is, Cathode side's oxygen flow rate, Pressure and Humidity) in the desired range provides the maximum throughput for the fuel cell. For analysis and identifying the parameter combination and range a neuro learning scheme is derived.

4 Discussion and Validation of Results

During PEM fuel cell operation, Hydrogen ions are collected in the electrolyte and electrons in the electrodes which actually represent a condition similar to charge accumulation in capacitor. As a result of this capacitive effect, the cell voltage cannot immediately follow the current variation in current. Models build by considering this double layer capacitive effects give more accuracy in representing the dynamics of fuel cell operation. Validation typically means to what extend the accuracy supports. In this work, a neural approach, is used to study the performance (get maximum cell power) and infer how variations in parameters affect the system. The experiment was run over the weka. associations platform with the help of java.

public java.lang.String toString()
double evaluateAttribute(int attribute) throws java.lang.Exception
toString in class java.lang.Object

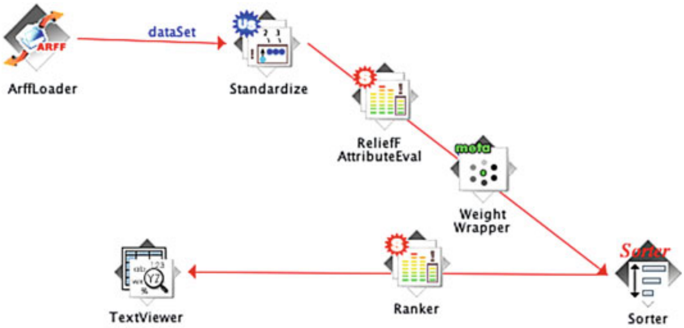


Fig. 5 Knowledge flow on Weka platform

For validating the result, dataset and the experimental setup used needs to be well explained. The experimental setup was run in four different environment, (a) without a controller (b) using Windrow Hoff alone (c) using Neuro Fuzzy alone (d) combination of windrow-hoff and neuro fuzzy. The parameters under consideration are pressure of oxygen, pressure of hydrogen, pressure of water, water injection rate, temperature, hydrogen flow rate, load and voltage. The test was run on 180 different combination of parameters. While running the test in Neuro-Fuzzy environment, 8 hidden layers were used. The layers were specifically fixed for abstracting maximum performance. The values of each parameter was falling in the range specified but for the performance to be maximum the correct combination is needed. This is what have been achieved in this process.

The database is in arff format, the dataset of the parameters are first standardised and the attributes are evaluated. The attributes which are not normalised are done a set of preprocessing and are made suited for prediction. A weight wrapper function is applied which makes the initial condition of the parameters to fo true in each iteration. The variation in each and every instance is noted and a sorter is applied over it to make sure the values are populating within a limit. The ranker then provides the accurate range in which the values need to be fixed for getting optimum performance. The text viewer displays the range in which parameters need to be fixed. Figure 5 shows the knowledge flow test setup.

Various parameter values were changed and the performance were calculated. The corresponding curve was also plotted automatically using the plot_graph function. We could see a clear overlap in the performance when the parameters are changed in each instance. Figure 6 shows the performance comparison for the controllers implemented and validation of optimum parameter set using weka. The straight line shows the performance of the system when no controllers are used. VI plot was seen to be 34:0 during the initial stage which changed to 23:20 during latter phase. The performance with WH alone and Neuro-fuzzy alone was also plotted and it was

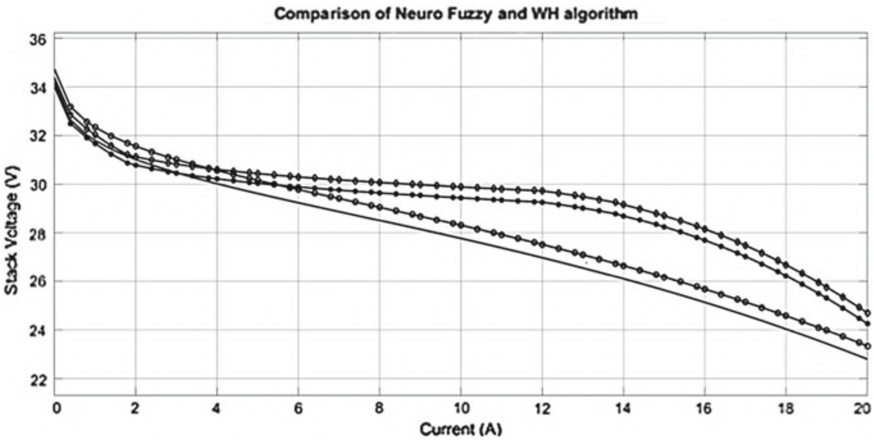


Fig. 6 Performance comparison of algorithm on stack voltage

seen to be having an improved performance than the no-controller environment. VI characteristics when using a combination of WH-NF was found to be 35:0 to 25:20. An improvement of 38% in performance was noted during each test cases analysis.

5 Conclusion and Future Scope

A control strategy was designed to achieve the following requirements (1) to maintain required value of cathode inlet pressure (P_c) by controlling the pressure from air compressor with respect to load current (I_{fc}), (2) to maintain required air flow (W_{co}) by controlling the motor according to load current (I_{fc}), (3) to maintain desired value of relative humidity (R_{hc}) by controlling the water injected (W_{inj}) into the humidifier for specific load conditions based on the value of dry air pressure and air flow from motor. Validations of the results were carried using weka platform. The results strictly adhere to the values obtained during modelling. A ranker mechanism was also used during preprocessing to prune the lower order parameters. As the neural net validates the model to be perfectly in align, it could be concluded as a perfect implementation of the strategy.

References

1. Lee, J.-Y., Cha, K.H., Lim, T.W., Hur, T.: Eco-efficiency of H2 and fuel cell buses. Int. J. Hydrog. Energy **36**, 1754–1765 (2011)
2. Hwang, J.J., Chang, W.R.: Life-cycle analysis of greenhouse gas emission and energy efficiency of hydrogen fuel cell scooters. Int. J. Hydrog. Energy **35**, 947–956 (2010)

3. Shaw, L., Pratt, J., Klebanoff, L., Johnson, T.: Analysis of H2 storage needs for early market "man-portable" fuel cell applications. Int. J. Hydrog. Energy **38**, 2810–2823 (2013)
4. Genconglu, M.T., Zehra, U.: Design of a PEM fuel cell system for residential application. Int. J. Hydrog. Energy **34**, 5242–5248 (2009)
5. Bernardi, D.M., Verbrugge, M.W.: A mathematical model of the solid- polymer-electrolyte fuel cell. J. Electrochem. Soc. **139**(9), 2477–2491 (1992)
6. Amphlett, J.C., Bumert, R.M., Mann, R.F.: Performance modelling of the ballard mark IV solid polymer electrolyte fuel cell II: empirical model development. J. Electrochem. Soc. **142**, 9–15 (1995)
7. Fowler, M.W., Mann, R.F., Amphlett, J.C.: Incorporation of voltage degradation into a generalised steady state electrochemical model for a PEM fuel cell. J. Power Sources **106**, 274–283 (2002)
8. Amphlett, J.C., Mann, R.F., Peppley, B.A., et al.: A model predicting transient responses of proton exchange membrane fuel cells. J. Power Sources **61**(1), 183–188 (1996)
9. El-Sharkh, M.Y., Rahman, A., Alam, M.S.: A dynamic model for a stand-alone PEM fuel cell power plant for residential applications. J. Power Sources **138**, 199–204 (2004)
10. Pukrushpan, J.T., Stefanopoulou, A.G., Peng, H.: Modelling and control for PEM fuel cell stack system. In: Proceedings of the 2002 American Control Conference, Anchorage, UK, USA, 8–10 May 2002, vol. 4, pp. 3117–3122 (2002)
11. Sharifi Asl, S.M., Rowshanzamir, S., Eikani, M.H.: Modelling and simulation of the steady-state and dynamic behaviour of a PEM fuel cell. Energy **35**, 1633–1646 (2010)
12. Xizhao, W., Chunru, D.: Improving generalisation of fuzzy if-then rules by maximising fuzzy entropy. IEEE Trans. Fuzzy Syst. **17**(3), 556–567 (2009)

The Systems' Integration Conception for Transactional Costs Reduction in Polish Financial Companies

Ewelina Wróbel and Marcin Hernes

Abstract Transactional costs are one of the key group of costs in business operations related to the provided goods or services. Reduction of these costs, while maintaining the quality of the product at the level expected on the market is the goal of monitoring and managing expenses in companies. One of the methods for reduction the transactional costs is an integration of IT systems. The aim of this paper is to develop a conception of integration of IT systems used by Polish financial institutions. Taking advantage of this conception in business practice may lead to a reduction of transactional costs.

Keywords IT systems' integration · Transactional costs · Reduction of costs

1 Introduction

Financial companies conduct permanent activities to acquire new clients, who are becoming more and more demanding and aware of their needs. Financial institutions must implement a high level IT tools in each area of the organization's operation to meet the expectations of the market. For this purpose ERP-class IT systems are used.

At the same time it is worth mentioning that integration covers both internal systems and their integration with external systems. This is a cost-effective solution, since each system may be integrated with several independent systems via an individually designed interface.

Financial institutions note multiple benefits of deviating from a complex approach (the main disadvantage is the organization's dependence on the one external or internal entity being the IT system provider) for the benefit of a diversification of responsibility and competences towards several providers.

E. Wróbel · M. Hernes (✉)
Wrocław University of Economics, ul. Komandorska 118/120, 53-345 Wrocław, Poland
e-mail: marcinhernes@ue.wroc.pl

E. Wróbel
e-mail: ewelina.wrobel@ue.wroc.pl

M. Huk et al. (eds.), *Intelligent Information and Database Systems: Recent Developments*, Studies in Computational Intelligence 830,
https://doi.org/10.1007/978-3-030-14132-5_32

However, the process of integrating the systems of various providers is not effective yet. This problem substantially extends the term of execution, Additionally, the effects of integration are often insufficient to support undertaking effective business decisions.

The aim of this paper is to develop a conception of integration of IT systems used by Polish financial institutions. The use of this conception in business practice may lead to a reduction of transactional costs. These are costs of business coordination resulting from the cooperation of many business entities on the market [1].

On the basis of a review of literature, observation of business phenomena, and based on professional experience, the authors of this paper have examined the relations between the level of use of IT systems integration tools in financial institutions in order to reduction of transactional costs, This is the main contribution in considered field.

The first part of this paper presents stat-of-the-art and basic notions in the analyzed area. Next, the systems' integration conception is presented. The discussion on anticipated results of integration is presented in the last part of paper.

2 Related Works and Basic Notions

Using IT systems in cost reduction is considered in many theoretical and practical works. For example [2] presents using RFID technology to cost reduction in public transport systems. Paper [3] presents method for finding an optimal choice of maintenance actions to be performed on a selected group of machines for manufacturing systems. The infrastructure needed to develop the Integrated Delivery Systems and the obstacles to its implementation in order to transactional cost reduction in healthcare institutions, is presented in [4]. The work [5] propose an integrated vendor–buyer inventory model in order to supply chain cost reduction. The work [6] states, in turn, that integration and standardization of IT systems in different organizations can lead to different costs reduction, including transactional costs. More often costs are reduced in financial institution by using different kind of cloud computing services [7]. Authors of the work [8] state, that using integrated IT systems is insufficient to cost reduction. It is also necessary to improve business processes for this purpose. The role of IT systems in financial institutions is still increasing. They support fundamental transformation of financial institutions in recent decades and point out the changes in the norms, regulations, and organizations [9, 10].

The cost of the company's operation comprises all components of assets (fixed and working capital), intangible assets, remuneration of employees and payments such as social insurance, used or paid for the purpose of producing goods or performance of services [11].

Cost management consists of, amongst others, their reduction during maintaining the quality of generated goods or services. Transactional costs of coordinating the company's operations result from collaboration of many entities on the market. The literature of subject presents several manners of cost classification. One of them

Table 1 Costs included in individual categories of transactional costs

Classification	Ex ante	Ex post
Costs included in the classification	– Costs of collecting market information (on clients, demand) – Costs of collecting information on competition – Costs of marketing and advertising – Costs of building the brand and consumer awareness of the brand and product – Legal and organizational costs of preparing the contract – Costs of negotiations	– Costs of monitoring and enforcement of contractual clauses – Costs resulting from improper assessment of contractual risk – Costs of resolution of disputes (legal, organizational, and costs of damages paid)

Source Own elaboration, based on [12]

Table 2 Examples of costs included to individual categories

Market TC	Managerial TC	Political TC
– Costs of the organization's operation on the market (marketing, building awareness) – Costs of collecting information about demand and supply – Costs of collecting information on client preferences – Costs of obtaining information on competition	– Costs arising within the enterprise	– Costs relating to the legal and business environment (e.g. costs of personal data protection, costs of safety of sensitive data)

Source Own elaboration, based on [14]

consists of setting a threshold contract execution point (e.g. signing a contract with a client) and dividing the costs into ex ante costs (incurred until the threshold point is reached) and ex post costs (all costs incurred afterwards) [12]. Classification of costs for the contract consisting in a sale of a financial product is presented in Table 1.

Another classification of transactional costs is proposed by E.G. Furubotne and R. Richter. They classified the costs into [13]:

- Market transactional costs,
- Managerial transactional costs,
- Political (public) transactional costs.

Examples of costs, which may be included in the above mentioned categories in a financial institution are presented in Table 2.

One of the organizations' needs related to reduction of transactional costs is an integration of IT systems. The next section of paper presents a conception of this integration.

3 Systems' Integration Conception

Financial institutions in their operation acquire various information about clients. Exemplary sources of such information include:

- Registration data,
- Information from business and credit databases,
- Data from business intelligence companies,
- Government registers,
- Data from anti-fraud registers,
- Data from falsification prevention registers.

The methods for acquiring data about clients by financial institutions include:

- Collection of data by the enterprise,
- Cooperation with institutions that collect data,
- Cooperation with institutions that aggregate data.

The volume of data stored by various IT systems used by companies does not allow for their effective sharing using non-automated information channels (e.g. via traditional email messages prepared and sent by system users) [15, 16]. Additionally the issue of ensuring data safety, intensified following the entry into force of the General Data Protection Regulation 2016/679 (GDRP)[1] results in that data processing and sharing should be performer in a structured manner, in accordance with the agreed communications between organizations. All events related to transferring personal data should be registered. Non-automated data sharing has one more defect, which causes an increasing interest of enterprises in IT systems integration, namely processing a large volume of data results in the need to invest significant human resources in the organization followed by significant costs.

In our conception, to limit the number of entities in the process of information management, we propose to use the information gates by financial institutions to obtain data from market. The data flow between entities and the gate is presented on Fig. 1.

The use of an information gate allows integrating the internal ERP or CRM with the gate, and, what follows, facilitate an asynchronous mechanism of sending client

[1]Pursuant to new regulations, personal data mean data that allow the identification of a natural person. This may include information such as: first name, surname, PESEL number, gender, email address, but also less obvious data such as the IP number, data on location, genetic code, political views or purchasing history. Any and all information collected about a given person, which allow identification, constitute personal data regardless of whether processed as paper copies or in digital form. https://www.gov.pl/cyfryzacja/rodo-informator.

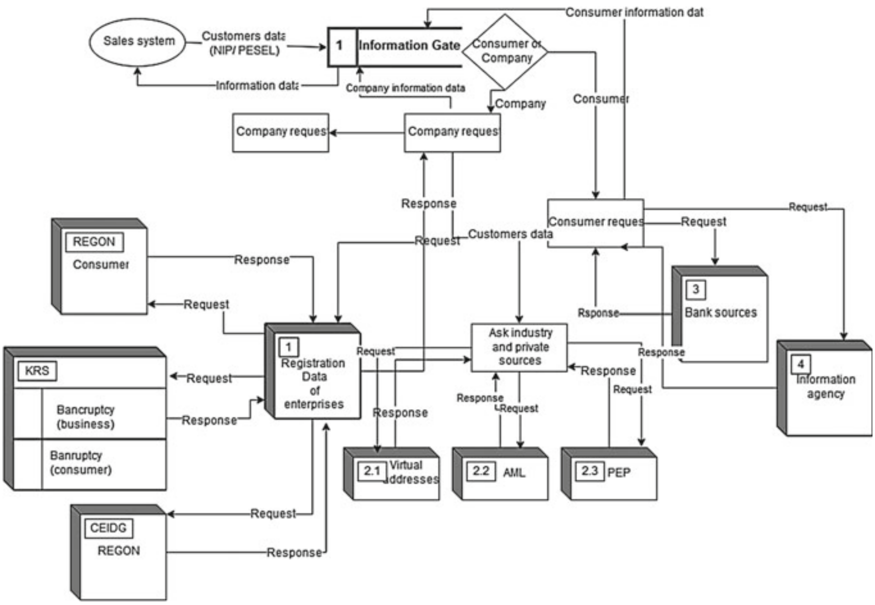


Fig. 1 Data flow between entities and information gate

inquiries to various sources at the same time. It is the gate's tasks, which acts as a proxy, to sent inquiries to sources and to send a response to the polling system. Integration with the gates usually occurs via the API provided by its administrators or a web service. The use of SSL, TLS and HTTPS protocols allows to ensuring the safety of transferred data.

Registration data of enterprises are acquired from: National Court Register (KRS), Central Registration And Information On Business (CEIDG) and Register of National Economy (REGON). Industry and private sources include: Virtual Addresses, Anti-Money Laundering[2] (AML) and Polish ePayment (PEP). Economics information sources include: National Register of Debts (KRD), Economic Information Bureau (ERIF) and National Economic Information Bureau (KBIG).

Companies offering financial products acquire market data from the following sources:

- Confederation,
- Collection of information on resignation from the offer,
- Market analysis,
- Monitoring existing investments,
- Products department,

[2]Set of procedures, laws and regulations designed to stop the practice of generating income through illegal actions.

No.	Company	VEHICLES	Passenger	Trucks	weight up to 3.5 tons	weight over 3.5 tons
1.	akf leasing Polska S.A.	0.5	0.5	0.0	0.0	0.0
2.	Alior Leasing Sp. z o. o.	1 275.2	633.2	270.2	208.3	61.9
3.	BNP Paribas Lease Group Sp. z o. o.*	685.2	518.6	142.4	0.0	142.4
4.	BZ WBK Leasing S.A.	1 327.7	689.9	212.6	106.1	106.5
5.	Caterpillar Financial Services Poland Sp z o	0.0	0.0	0.0	0.0	0.0
6.	De Lage Landen Leasing Polska S.A.	193.8	4.0	11.8	11.8	0.0
7.	Deutsche Leasing Polska S.A.	19.2	1.6	0.0	0.0	0.0
8.	Europejski Fundusz Leasingowy S.A.	1 811.9	885.9	454.0	260.8	193.2
9.	FCA Leasing Polska Sp. z o. o.	196.7	110.5	86.3	86.3	0.0
10.	Grupa Idea Getin Leasing	3 137.4	1 462.3	579.2	375.5	203.7
11.	Grupa Masterlease **	484.7	448.0	27.4	26.8	0.6
12.	Impuls-Leasing Polska Sp. z o. o.	340.9	176.9	52.3	35.7	16.6
13.	ING Lease (Polska) Sp. z o. o.	1 270.5	469.8	198.5	114.4	84.2
14.	Leasing Polski Sp. z o. o.	56.6	7.8	12.2	5.5	6.7
15.	MAN Financial Services Poland	289.6	0.0	18.0	1.9	16.1
16.	Millennium Leasing Sp. z o. o.	1 030.4	407.3	182.4	94.8	87.6
17.	mLeasing Sp. z o. o.	1 947.2	1 298.2	261.7	152.8	108.9
18.	NL-Leasing Polska Sp. z o. o.	28.4	4.5	5.3	4.2	1.2
19.	PEAC (Poland) Sp. z o.o.***	10.6	5.9	0.5	0.2	0.2
20.	Pekao Leasing Sp. z o. o.	1 044.2	396.2	141.3	53.9	87.4
21.	PKO Leasing S.A.	3 037.3	1 777.2	344.8	154.5	190.3
22.	PSA Finance Polska sp. z o.o.	421.5	327.8	93.7	93.7	0.0
23.	Santander Consumer Multirent Sp. z o. o.	480.9	434.7	40.9	40.9	0.0
24.	Scania Finance Polska Sp. z o. o.	482.5	1.4	51.2	0.0	51.2
25.	SG Equipment Leasing Polska Sp. z o. o.	386.4	73.8	47.8	14.4	33.4
26.	SGB Leasing Sp. z o. o.	21.6	11.1	5.5	5.1	0.4
27.	Siemens Finance Sp. z o. o.	16.1	0.6	7.8	0.2	7.6

Fig. 2 Sample report on results in the leasing sector prepared by The Polish Leasing Association (ZPL). *Source* [15]

- Quarterly reports (information about competition—an exemplary report on financial results in the leasing sector is presented in Fig. 2).

Each data may be used and processed to a various degree. Examples of storing and using such data include:

- Collection of data from external sources in real time, as needed. This may result in a reduction of costs relating to maintenance of HDD required to store data, and to a reduction of expenses on data protection.
- Storage of processed or unprocessed data in an internal database (having regard to validity of inquiry).
- Data processing via transaction and entity assessment cards (scoring, rating).
- Processing via mapping, i.e. assignment of a mapping value to each mapped piece of data. Data mapping examples are provided in Table 3.
- Decision engines.

Table 3 Mapping the data on historical transactions of a business entity to assess its reliability

Days of delays in repaying the debt	Customer's quality assessment
0	Very solid
1	Reliable
2–4	Good
5–10	Average
11–30	Risky
31–100	Unreliable
>100	Blocking offers

The gates uses such methods of IT systems integration as:

1. Exchange of information using Web Services. This is a standard, which structures and puts in order the manner of data exchange between systems using the network [16]. The following protocols are used:
   - Simple Object Access Protocol (SOAP) for communication—used to transmit remote calls,
   - Web Services Description Language (WSDL)—used to distribute network connection parameters base specification,
   - Universal Description, Discovery and Integration (UDDI)—used to register of shared components in service.
   - Representational State Transfer (REST)—defines a set of constraints used for creating web services.
   - Extensible Markup Language (XML)—used to exchange the data in uniform standard.
2. Application programming interface (API)—includes specifications for data structures, routines, object classes, variables and remote calls. The client sending an inquiry for a defined action with parameters. The server accepts the requests, processes it, generates a response and sends to the client.

In opposite to our approach, the data warehouse can be used in order to data integration and information processing. Extract, Transform and Load (ETL) and Online Analytical Processing (OLAP) can be very useful in this purpose. However, the costs of developing and maintenance of such warehouse are higher than costs of the gate (the large amount of historical data has to be stored in such warehouse). In the gate approach, these data are stored in financial entities' databases and data warehouses, and only data required in given moment are received by financial companies.

4 Transactional Cost Reduction—Discussion

The authors will take into consideration the classification of transactional costs described in Table 1, to analyze transactional cost reduction by IT system integration. Each of these groups generates different costs and they should be not be analyzed collectively.

4.1 *Reduction of Ex Ante Costs*

Ex ante costs are the costs generated until the transactional threshold point (in a financial institution, this is the moment of signing the contract). Costs, which may be reduced using the IT systems integration include:

1. Marketing costs. IT systems integration allows to reducing marketing costs through using the RSS (Really Simple Syndication) [17] meaning an XML-based format, serving for transferring information headings. Content maintenance in several different locations is connected with investing time into editing and administration. RSS channels allow automatic updates of inquiries to an external website once the content is modified in any manner. As a result, customers have always actual information related to financial company's offer.
   Another method to reduce marketing costs is processing the data about the clients of the financial institution. The systems managed by marketing departments may use financial and accounting data collected in another system. The prepared advertising campaigns may be properly profiled and reach the defined group of recipients, in consequence. This also allows to a better adjustment of advertising content addressed to clients, reducing costs incurred for advertising which turns out ineffective.
2. Costs of collecting a client data. Client data are acquired from various sources. Some of them (such as first name and surname, company name, addresses) may be acquired from clients directly. However, this extends the time dedicated to completing necessary documents. Currently, clients expect a financial institution to shorten the time of handling applications to a minimum. One of the solutions allowing to limit the time dedicated by an adviser to complete the application is to take advantage of integration with the gates which sent inquiries to registration sources (databases such as the KRS or REGON). This allows to quickly acquiring a client data (such as company name, address details, registration details, contact details and details of company owners) based on the provided tax identification number (NIP). The benefits of this solution include [18]:

   - Reduction of costs dedicated to complete several dozen boxes on a financing request,
   - Reduction of costs resulting from errors during data entering,

- Reduction of costs incurred to verify data provided by the client (in case of integration with the gates, data is collected from government registers and additional verification is not required),
- Shortening the time needed to prepare the request and, what follows, minimizing the risk of losing clients being in the group of people feeling discouraged by the time-consuming nature of completing formalities.

3. Costs of acquiring historical data on transactions and client assessments. Some data on clients are difficult to acquire, others are impossible to acquire (for example, data related to the credit history of the potential clients). Such data may be acquired from business or credit information databases. Currently, the main source of business and credit information in Poland is the BIK group, which includes Biuro Informacji Kredytowej S.A. and Biuro Informacji Gospodarczej InfoMonitor SA. This group processes information obtained from over seven hundred financial institutions [19]. Since this is an institution established on the basis of the Banking Law,[3] it is the only one in Poland allowed to collect and process such data legally. Data possible to acquire from the aforementioned databases, which affect the reduction of transactional costs include:

   - Personal data,
   - Liabilities,
   - Debt,
   - Bank accounts,
   - Data on debtors (consumers and companies).

   For a financial institution, data obtained from this source constitute a source of valuable knowledge and allow to minimizing the costs related with preventing fraud or conclusion of contracts with insolvent clients.
4. Costs of transactional risk assessment. The costs of assessing the transactional risk constitute a particularly sensitive area of conducting business activity. Reduction of costs in this area may be related with significant risk in connection with expenses to be incurred for debt collection activities. Some of the receivables are uncollectable, hence the costs of an error in the area of this risk may result in losses in the amount equal to the amount of entire transactional costs. Method to secure the reduction of costs in the risk area are automation and optimization processes. Risk assessment for a transaction may be conducted in an external system, which allows generating client and transaction risk assessment cards (rating cards, scoring cards). Their integration with sales systems, systems used by AML departments, anti-fraud departments, BIK systems and post-sales sys-

[3] Article 105.4, Banking Law—banks, together with banking industry chambers, may establish institutions to collect, process and transfer: (1) to banks—information constituting bank secret, if such information is needed in connection with the performance of banking activities; (2) to other institutions with statutory authorization to grant credits—information about claims, turnover and balances on bank accounts, if such information is required in connection with granting credits, loans, bank guarantees and warranties. http://prawo.sejm.gov.pl/isap.nsf/download.xsp/WDU19971400939/T/D19970939L.pdf.

tems allows collecting the processing data from various sources in a quick and automatic manner. Therefore, the risk related to errors occurring in a conception, where every transaction is assessed by a risk analyst, is limited. This pattern is also less vulnerable to possible frauds, thus transactions are less risky.

4.2 Reduction of Ex Post Costs

1. Costs of post-sales handling. Post-sales handling of a contract is a process commencing upon the activation of a contract regarding a financial transaction. At this stage, transactional costs incurred by the organization include costs of monitoring and recording client payments and receivables. To automate the process, integration may be performed with:
   - The bank system, which allows recording and monitoring payments in real time,
   - The mail server, which allows sending messages on the status of payments from clients directly from the application, using automatic triggers,
   - The sales system, which allows blocking transactions of insolvent clients,
   - The transactional risk assessment system, which allows obtaining the BIK classification and prevent risky transactions.
2. Reporting costs are costs incurred by the organization as a result of handling ordered manual reports. These reports may be generated automatically, and their scope may be defined by the certain parameters. With the use of an appropriate Graphical User Interface (GUI), where the user may define the parameters of the request sent to API, reports may be configured in an ongoing manner, without the need to engage additional funds.

5 Conclusions

The integration of IT systems used in financial institutions is not only an opportunity to ensure safety related to becoming independent from one provider. As presented above, it is an opportunity to reduce transactional costs incurred by the organization in its daily operation.

Costs may be reduced in two ways, one of them is a make change of the process, and the second one is process improvement and optimization. Authors proposed a solution appropriate for mature organizations, with high process quality, who want to take advantage of the opportunities granted by technologies allowing the integration of IT systems, to improve their processes and ensure that they are more effective, efficient and, thus, less costly. This solution can be used by financial companies in practical implementations.

The main limitation of the presented approach is lack of uniform, standardized methods for processing information acquired by gate. These processes are realized separately by IT systems running in particular financial companies. Therefore, in case the same information, conclusions in particular organizations may differ.

The further research works may be related, among other, to implementation of proposed solutions in real IT system running in financial institution and to analyze the performance of the results of integration.

References

1. Williamson O.E.: Transaction-cost economics: the governance of contractual relations. J. Law. Econ. **22**(2) (1979)
2. Chou, S.Y., Ekawati, Y.: Cost reduction of public transportation systems with information visibility enabled by RFID technology. In: Chou, S.Y., Trappey, A., Pokojski, J., Smith, S. (eds.) Global Perspective for Competitive Enterprise. Economy and Ecology. Advanced Concurrent Engineering. Springer, London (2009)
3. Liu, F., Zhu, H., Liu, B.J.: Cent. South Univ. 20: 1559. https://doi.org/10.1007/s11771-013-1648-y, 2013
4. Al-Saddique, A.: Integrated delivery systems (IDSs) as a means of reducing costs and improving healthcare delivery. J. Healthc. Commun. **3**(1), 19 (2018)
5. Shahpouri, S., Fattahi, P., Arkan, A. et al.: Int. J. Adv. Manuf. Technol. **65**, 657. https://doi.org/10.1007/s00170-012-4205-1 (2013)
6. IT Optimization—reducing costs without diminishing returns. In: Buchta, D., Eul, M., Schulte-Croonenberg, H. (eds.) Strategic IT Management. Gabler (2007)
7. Benjamin-Addy, E.L: The Banking Sector in Ghana: Issues in Relation to Current Reforms. Anchor Academic Publishing (2013)
8. Nicoletti, B.: Cloud computing in financial institutions. In: Cloud Computing in Financial Services. Palgrave Macmillan Studies in Banking and Financial Institutions. Palgrave Macmillan, London (2013)
9. Nicoletti, B.: Financial institutions and information and communication technology. In: Cloud Computing in Financial Services. Palgrave Macmillan Studies in Banking and Financial Institutions. Palgrave Macmillan, London (2013)
10. Ćetković, J., Lakić, S., Knežević, M., Žarković, M., Sazonova, T.: The use of transaction costs theory in interorganizational design. MATEC. Web. Conf. **53**, 01055 (2016). https://doi.org/10.1051/matecconf/20165301055
11. Hansen Henten, A., Maria Windekilde, I.: Transaction costs and the sharing economy. Info. **18**(1), 1–15 (2016)
12. Hannon, M.J., Foxon, T.J., Gale, W.F.: 'Demand pull' government policies to support product-service system activity: the case of Energy Service Companies (ESCos) in the UK. J. Clean. Prod. **108**, 900–915 (2015)
13. John, G., Reve, T.: Transaction cost analysis in marketing: looking back, moving forward. J. Retail. **86**(3), 248–256 (2010)
14. Furubotn, E.G., Richter, R.: Institutions and Economic Theory. The Contribution of the New Institutional Economics. The University of Michigan Press, Ann Arbor, Michigan (2007)
15. http://www.leasing.org.pl/pl/statystyki/2018 (retrieved: 20.09.2018)
16. Pielecka, M.: Integration of information systems—interorganizational information exchange, http://zif.wzr.pl/pim/2013_1_4_28.pdf (in Polish) (retrieved: 20.09.2018)

17. Chaturvedi, A.: Automated Web Service Change Management AWSCM—A Tool. In: 2014 IEEE 6th International Conference on Cloud Computing Technology and Science, p. 715. (2014) https://doi.org/10.1109/cloudcom.2014.144. ISBN 978-1-4799-4093-6
18. https://www.inflibnet.ac.in/caliber2009/CaliberPDF/42.pdf (retrieved: 20.09.2018)
19. https://www.bik.pl/o-nas (retrieved: 20.09.2018)

Design of Reactive Systems for Control Network Traffic on the Kubernetes Platform

Lubos Mercl, Vladimir Sobeslav and Peter Mikulecky

Abstract The container virtualization on the Kubernetes platform brings challenges that need to be addressed when a production or test load is running across the cluster. If applications running in containers are spread across the Kubernetes cluster, resource usage may be inefficient across the environment, which may result in overloading of individual nodes and inefficient load on others. One area where inefficiencies may occur is the load on network lines and network communications. This article discusses two algorithms and approaches that can be applied to the Kubernetes platform while helping to manage network traffic and lines across the cluster, which can make the cluster components more efficient. Both algorithms collect the monitored data from the cluster, but each one the data collected behaves differently, and data transformation and analysis takes place in another part of the system. The first algorithm is an agent-based algorithm that collects and performs basic data analysis and is capable of responding to detected information. The second is the algorithm that collects the data into the central element of the system and then analyzes it and, based on the information obtained, controls the individual components of the cluster.

Keywords Container virtualization · Kubernetes · Network traffic analysis · Algorithm · Complexity

1 Introduction

The traditional virtualization brought primarily a release from physical resources to virtual resources when the whole concept was based on virtualization of the virtual servers running the individual applications that were being used.

The new approach to development and developing applications and services is the container virtualization that runs container applications, which are separate portions

L. Mercl (✉) · V. Sobeslav · P. Mikulecky
Faculty of Informatics and Management, University of Hradec Kralove,
Rokitanskeho 62, 50003 Hradec Kralove, Czech Republic
e-mail: lubos.mercl@uhk.cz

M. Huk et al. (eds.), *Intelligent Information and Database Systems: Recent Developments*, Studies in Computational Intelligence 830,
https://doi.org/10.1007/978-3-030-14132-5_33

of the operating system where these services run. The container virtualization brings benefits primarily to the continuing less demanding system resources, which is even smaller than traditional virtualization. A further advantage is the lower technical difficulty and run only over one core of the operating system, making system resources of one operating system running in one physical or virtual operating system.

However, the container virtualization also has drawbacks that may be limiting for individual deployments and solutions built on container virtualization. First of all, this is not true virtualization, because the containers share the same operating system kernel and the same operating system.

In large environments, the container virtualization becomes to be very complex and management can be complicated, because there is huge amount of container cluster nodes, which perform an platform where individual containers run. On these container clusters, there is needed some management tool called the orchestrator, which can manage whole cluster and individual instruments. Among the used orchestrators belongs the Docker [5], Docker Compose, Docker Swarm, Apache Mesos [1], or Kubernetes [13]. Kubernetes is the most comprehensive, offers the greatest opportunities for container cluster orchestration and possible custom extensibility, and is concerned with this work [16].

2 Problem Definition

For the container virtualization and generally virtualization, essentially cloud computing, it is typical that the end user or administrator does not have much eminent interest in where his application or service is running. It is only interested in the fact that the service is running and if it meets the performance parameters, e.g. it is in line with the service-level agreement.

Service level is often defined simply as service availability. Some cases, however, further develop this level of disintegration into individual service parameters, such as a response to a database application or the application itself.

The application's response can then be divided into simple attributes and parameters that are important to the application speed. Such attributes may be the correct application architecture, resource allocation, prioritization to shared resources, or the right network architecture design, a sufficiently permeable network, and other network parameters are properly optimized.

2.1 Network Inefficiencies

Clusters, where the container virtualization runs, can offer a lot of space where inefficient cluster components can be loaded. For example, the simplified Kubernetes cluster consists of several nodes that are networked. One or more nodes are so-called master nodes that control the entire cluster, and the remainder are so-called nodes or minions running on individual containers.

Already from this simplified cluster, it is possible to identify possible potential areas that may be the reason or the point of inefficiency. These include:

- High load on individual lines between clusters,
- Load specific network ports on nodes,
- API load on any master node,
- The difference between the loads of individual work nodes,
- Large network overhead on one of the work nodes,
- Running applications that communicate directly across multiple work nodes on the network and thus unnecessarily load network lines, etc.

To solve such problems, it is possible to follow the knowledge and best thoughts of those who operate and develop such environments, but it is important to minimally monitor trends in such environments using a load tracking system, which is still desirable to complement an environment intelligence system in such environments, which would evaluate each of the monitored attributes, searched in context data and responded to current and predictive load.

It is possible to say with certainty that the larger the cluster, the higher the level of entropy that can affect the whole cluster. The same applies in the case of the complexity and complexity of the user's workload of the cluster, the higher the level of entropy, and the need to eliminate the inefficiency and to make the cluster more efficient.

2.2 Attributes and Important Features of Network Load

For the proper functioning and definition of such a system for ambient intelligence and the elimination and resolution of cluster entropy, it is primarily necessary to define metrics, attributes, and system components that need to be tracked. This definition may be different for each cluster, but in many respects it will always be similar, and similar or similar metrics will be followed, with only a different view and evaluation [11].

The following Table 1 shows, for illustration, the basic attributes that can be observed in the Kubernetes cluster and which are important. The table also captures where these attributes can be tracked and the data are collected, as well as the units that can acquire these attributes.

The enumeration of these attributes is illustrative only, and these are basic attributes that can be tracked, but other attributes can be tracked using other parameters and add-ons, which are appropriate for tracking and deployment.

Generally speaking, however, it is important and possible to observe the 5 basic attributes in the Kubernetes cluster:

Table 1 Network attributes in Kubernetes clusters

Attribute	Location	Value
Network line load	Node in the cluster	Data transferred per time unit
Network port load	Node in the cluster	Data transferred per time unit
Connection quantity	Network interface or port	Number of connections
Connection map	Master node	Connection map between nodes
API queries quantity	Node in the cluster	Number of API queries per time unit on in time frame
Container resources utilization	Node in the cluster	Using system resources over time
Metrics used by applications	Node in the cluster	Optional metrics that are relevant and important to tracking
Containers quantity	Node in the cluster	Number of containers per node

- The network infrastructure, which includes the first network and infrastructure services,
- Physical nodes,
- Applications or pods and deployment,
- Containers,
- Kubernetes orchestration and its components.

3 Related Works and Existing Solutions

This work is based on analysis of container virtualization orchestrator and their possibilities [16], cloud security analysis work [9], which defines attributes and areas, which should be handled in security solution on cloud platform. Also work takes inputs from Ambient Intelligence solution for cloud environments [11].

The options that exist for network communication analysis and Kubernetes cluster tracking can be divided into several parts.

The first part is the monitoring tools built in Kubernetes, or the tools developed for Kubernetes clusters and belonging to open source tools or their development by a community or company. This group includes, for example, Heapster [12], which is below, which is installed inside Kubernetes and collects data and records from logs from containers and pods. Heapster can be complemented by the Grafana project [7], which offers visualization of collected data, as well as the InfluxDB project [10], which offers the possibility of a highly available database for collected data. Prometheus [21], an open source project offering options for data collection, visualization and alerting, is also available for use. The CAdvisor project [6] focuses on performance analysis of containers in Kubernetes clusters. Other tools are New Relic [19], Weave Scope [28], Sysdig [25], and Datadog [3].

Another group are tools that can also be used for Kubernetes, but they can predominantly track only basic network parameters without container virtualization superstructure. These include, for example, PRTG Network Monitor [22], which is more about network infrastructure. Other tools include Nagios [18], Zabbix [29], OpenNMS [20], Logic Monitor [15] and more.

It is also possible to omit from the analysis also scientific papers that deal with the given issue. Many of them work deals with network attribute analysis systems without linking to container virtualization. Between interesting works belong, for example, Depren [4], which deals with an intelligent anomaly detection system in computer networks, or Roshan [23], which deals with a network intrusion detection system using learning machines.

Further work is already being based on cloud infrastructures and cloud computing. These include, for example, Chi [2], which deals with a cloud platform design for detecting and preventing security of a software-defined network. Grossmann [8] deals with the monitoring of container services in terms of network parameters. Tsai [26] presents a distributed analytic system on the Kubernetes platform and Sohal [24] addresses the framework for identifying security threats in cloud infrastructures. Vieira [27] and Zhang [30] deal with intrusion detection systems in Smart Grid networks. Modi [17] presents capabilities and analysis of basic techniques for detecting breakthroughs in computer cloud infrastructures. Lin [14] deals with the algorithm for selection and decision-making rules for detecting penetrations into the network and its applications. Cloud usage management and analysis is also important for security purposes and can help with whole environment security [9].

4 Proposed Solutions

A reactive system that collects data will analyze and then respond to the knowledge gained from the data. It has to be designed so that it can react to changes in the state at the right moment, and especially precisely.

The system should collect consistent data that will be stored for future analysis and statistical analysis of time series and search for correlations over time.

On the Kubernetes platform, individual nodes can be understood as nodes among which they are from the point of view of the edge network that can acquire values that are proportional to the transmission speeds of a given line or property of a given connection, so for the need of analysis, it is possible to use the rated edges to visualize and optimize the network and the cluster.

The system offers and supports the following options:

- Optimizing the use of lines and services and containers,
- Optimizing the use of individual nodes,
- Detection of anomalies due to known standard cluster behavior,
- Advanced planning of new services with past and current knowledge trends and the possibility of predicting other trends.

Any change and condition may or may not cause a change in cluster behavior and load, or trigger a Kubernetes response or cluster monitoring. Such changes may include, for example, that the use of several inter-communications services on multiple cluster nodes may result in higher usage of network lines that are subsequently loaded. This change can then trigger a reaction that mutually communicating services can be shifted to the same nodes when they do not burden network lines between nodes.

4.1 Basic System Architecture

The proposed system should have modules, which will cooperate to meet the requirements of users and also work optimally. These modules are Management console, Data collection module, Data processing module, Data storing module and Data visualization module.

Management console This module takes care about management of the whole system. It serves user interface, where users can modify some parameters, which are configurable to optimize the system with respect to environment specific or current needs, which may change from time to time.

The management console should be linked with the data visualization module and both module can be run on one front-end web page.

Data collection module This module collects data on agents. Depends on usage and collected data, there are a few ways, how can be implemented:

- Package, which is installed on node,
- Containers, which is run on the node and in this container, there is written some application, which collect data,
- Application part, which is written like part of deployed application,
- and others.

Implementation, how will be data collection module highly depends on specific requirements, environment or deployment, and there should be standardized message structure, which can be used on the most of type of specific possibilities.

Data processing module Between agents and other highest level of system, there should be implemented module, which will help with data processing and data transferring. This part of solution is important, because helps with optimization and acceleration of the whole system. It is required to respond as quickly and as correctly as possible to the detected information.

Thanks to specific optimization and implementation a few techniques can be reaction as faster as it is needed, but there can be also requirement to be not faster, but to data transfer fingerprint can be to smaller and only needed information to be transferred. This will require different technique from the requirement as faster as possible way. For example, Kubernetes cluster, which is run in data center, ill not need much transfer and fingerprint optimization as Kubernetes cluster, which is run like Internet of Thing solution, where each node is very distanced and it is important to save the transferred data.

Data storing module For keeping data and important values, there should be implemented some database, which will be store data. This module is very close to data processing module, which should prepare data for storing, and also with data visualization module, which read data from this database.

There can be implemented also two type of data store locations:

1. Short time fast store location, which is used mainly for data, which are needed quickly and often (e.g. for visualization).
2. Long time slow store location, which is used for all the data, which should be kept for long time, but there are needed only time from time or are not regularly needed, but there is request to have it.

There are a few ways and techniques, which can solve issues with requirements and also be helpful with possibilities for Big Data analysis.

Data visualization module This module provides the visualization of metric and data, which are collected and can be useful for environment checking and help with troubleshooting or management. There can be provided graphs of metrics with time, which can be simply modify to get different views for getting knowledge from these information. Also there should be some connection maps, current static values of environment and others.

The data visualization module should be complemented by basic statistical analysis, which will provide some basic calculations over detected and stored data. But these statistical analysis should included only simple calculation, which are not too demanding to computing.

Communication between all module depends on used technologies (e.g. used database engine, used visualization tools, programming language).

4.2 Data Processing System

But the question remains how this system should be properly designed architecturally. This can be divided into two possible solutions that will perform support for the central monitoring system, collect, transform, and then respond to data. These systems are:

1. An agent-based system where the agent collects data then performs basic analysis and transformations on the agent and passes it to the central system. After the agent's analysis, he or she may already react to change or knowledge. The central system then takes into account the entire cluster and can make changes at the cluster level.

2. A system based on a central and gateway system where all data is collected from individual cluster elements and sent to a central system that, if it finds room for optimization or reaction, does it here.

Both systems eventually store the data in a central database where data is stored in time series. This database is linked to a central system where the data is finally transformed and ultimately sent to the database.

Algorithm parameters for both solutions that affect the complexity of the algorithm are as follows:

- The number of nodes in the cluster,
- The number of cloud services and containers,
- The number of values being tracked,
- The ratio of control and control elements, controlling for example a central system, an agent in an agent system, or an element between a central system and an agent that analyzes the connected parts of the subset of the cluster and acts as a data-gathering agent to respond to them and forward them to a central system. This is, therefore, a gateway between the lower subset of the cluster and the higher part of the cluster. The controlled element is one where the data is mainly detected. Algorithm parameters for both solutions that affect the complexity of the algorithm are as follows:

5 Agents Based System and Its Algorithm

The first system is a system based on agents that collect data, care for their transformation, and at the same time realize basic analysis and reactive behavior on the level of information found at the level of the agent, thus streamlining the response speed to cluster changes.

Agents are then connected to a central element, or gateways can be implemented that can divide agents into several subsets as needed and perform data preparation, analysis, and reactive behavior on a lower set of agents, or they can help with data processing and preparation, thereby increasing the efficiency of processing and its speed. The central element then takes care of the entire cluster and controls the intelligence of the whole environment.

An illustration of such a system can be seen in Fig. 1, where one central element and several agents are displayed that communicate with the central element directly.

Since intelligence can be realized on virtually all elements, the reactivity to changes can be very fast, which can be a source of problems and non-deterministic behavior of the cluster, which must be treated in a real implementation by mechanisms that prevent unwanted behavior of the cluster.

Algorithm of this solution can be simply described like:

1. The data are collected on the agent, transformed, evaluated, and then sent to the higher element.

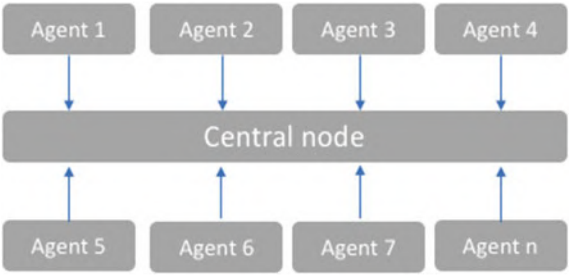


Fig. 1 Agents based system. *Source* author

2. The data are merged from multiple sources on the gateway and then sends it to a central element within one run.
3. The central node receives, transforms, evaluates, and saves data in the database.

Input values, which are influencing algorithm complexity are agent count (number of managed elements), value count (number of managed values) and central element count.

If there were more members of a central node when the communication comes in between multiple elements, the total number of steps would be divided by the number of these members.

6 Gateway System Based on the Central System and Its Algorithm

The second system is a system based on a central system or gateways that collects data from all the monitored elements and then analyzes and performs reactive actions.

The simplified architecture of the solution shown in Fig. 2, where there is one central component, several support nodes (gateway 1 to b) and to each gateway may be connected 1 to n agents.

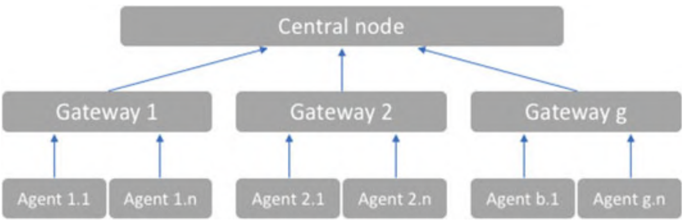


Fig. 2 Gateway system based on the central system. *Source* author

The central element takes care of the cluster run and analyzes the entire cluster and realizes ambient intelligence in the environment and, based on the defined and learned rules, it subsequently optimizes the environment and streamlines the use of cluster resources.

Implementation of a central element can be implemented if necessary by more central elements, among which the load and processing will be divided by the implementation of the chosen load distribution algorithm.

Algorithm of this solution can be simply described like:

1. The data are collected on the agent, transformed, evaluated, and then sent to the higher element.
2. The central node receives, transforms, evaluates, and saves data in the database.

Input values, which are influencing algorithm complexity are agent count (number of managed elements), value count (number of managed values) gateway count (number of used gateway role elements), layer count (number of layers in architecture) and central element count.

Also in this solution, there can be more central nodes so this can be helpful with processing data on second step.

7 Complexity Aspects and Influence of Transformation and Evaluation in both Data Processing Systems

For both systems, calculations of the number of steps are considered as a better solution, but it is not possible to forget the impact of transformation and evaluation on agents, which can burden the system unambiguously and the resulting complexity, especially time and computation, increase.

This influence is mainly due to the complexity of the data on their transformation and the specific implementation, which must be taken into account. For better and faster transformation, after analyzing the impacts in the implementation, it is advisable to consider increasing the resources at the places where the transformation takes place and thus optimizing the steps.

The complexity is for both solution can be reduced by distributed calculation on more elements. Lower layer elements can prepare data in suitable for highest elements in platform. This can reduce amount of calculated data on elements, where data from multiple elements are clustered. For these designs, the better architecture for this approach is solution with gateway usage, because this solution has more layer of architecture.

For the needs of the work, the transformation gate is one step, however, despite the various implementation and analysis required, it is possible to divide the transformation into the following steps:

1. Data retrieving,
2. Transformation steps for individual data or values,

3. Input data analysis (can be done before transformation),
4. Storing transformed data,
5. Sending transformed data,
6. Submission of evaluated responses to individual inputs.

These steps are only general and may vary according to the implementation needs and environmental differences.

8 Conclusion

This work deals with the analysis of the possibilities of monitoring the network parameters on the container virtualization Docker and especially with the use of the Kubernetes orchestrator, which is used to create, manage and control the container virtualization cluster.

The work is conceived in several units, the first is the presentation and analysis of container virtualization issues, the second part is the analysis of necessity and the possibility of monitoring the network parameters in the Kubernetes cluster and the third part is the analysis and presentation of algorithms and approaches for analyzing parameters and attributes.

The first system and algorithm is an agent-based algorithm that analyzes individual components on the node where they are located, performs basic data analysis and responds to data, transformed data then sends to a central element that performs a higher phase of the entire environment analysis, responds to the detected store data and data in a database.

The second system is a system based on a central element or gateways that collects data from individual elements, these data then transform and analyze and perform individual changes throughout the cluster.

Both algorithms have the same purpose but there are different approaches and usages and differences are mainly in the complexity of the individual clusters and in the necessity of monitoring the individual algorithms.

The agent system is more suitable to more complex environments and clusters where it is necessary to monitor and respond to higher amounts of data and changes and instead the gateway system fits into smaller or more generic and less flexible clusters.

The individual implementation of this architecture should be included some complexity analysis, which can resolved future issues with data processing. The architecture describes two approaches, which each one has some pros and cons.

9 Future Work

Further development of the algorithms and optimization and generalization of these algorithms, which can be applied to multiple systems with simple modification.

Also, statistical and predictive analysis should be optimized to automatically look for similarities in time series that should be as accurate as possible and as quickly and efficiently as possible. Both systems will also be good to maintain and develop concurrently because once it can be a suitable agent-based system and sometimes one central system may be more appropriate.

In any case, the issue is suitable for further investigation and analysis and, above all, it will be appropriate not only to conduct theoretical research and design, but also to make real deployment and implementation in real-world and real-time data. Possible research could be complemented by agent systems and behavioral simulations in the laboratory environment.

Acknowledgements This work and the contribution were supported by a Specific Research Project, Faculty of Informatics and Management, University of Hradec Kralove, Czech Republic.

References

1. Apache Mesos. http://mesos.apache.org/. Last Accessed 2 July 2018
2. Chi, Y., Jiang, T., Li, X., Gao, C.: Design and implementation of cloud platform intrusion prevention system based on SDN. In: 2017 IEEE 2nd International Conference on Big Data Analysis (ICBDA), pp. 847–852. IEEE, Beijing, China (2017)
3. Datadog. https://www.datadoghq.com/. Last Accessed 2 July 2018
4. Depren, O., Topallar, M., Anarim, E., Ciliz, M.: An intelligent intrusion detection system (IDS) for anomaly and misuse detection in computer networks. Exp. Syst. Appl. **29**(4), 713–722 (2005)
5. Docker - build, ship, and run any app, anywhere. https://www.docker.com/. Last Accessed 2 July 2018
6. Google/cadvisor: analyzes resource usage and performance characteristics of running containers. https://github.com/google/cadvisor. Last Accessed 2 July 2018
7. Grafana - the open platform for analytics and monitoring. https://grafana.com/. Last Accessed 2 July 2018
8. Grossmann, M., Klug, C.: Monitoring container services at the network edge. In: 2017 Proceedings of the 29th International Teletraffic Congress (ITC 29), vol. 1, pp. 130–133. IEEE, Genoa, Italy (2017)
9. Horalek, J., Sobeslav, V.: Analysis of the use of system resources for cloud data security. In: 15th International Conference on Mobile Web and Intelligent Information Systems, MobiWis 2018, Barcelona. Lecture Notes in Computer Science, vol. 10995, pp 101–111. Springer, Berlin (2018)
10. InfluxData (InfluxDB) - time series database monitoring and analytics. https://www.influxdata.com/. Last Accessed 2 July 2018
11. Komarek, A., Pavlik, J., Mercl, L., Sobeslav, V.: Hardware layer of ambient intelligence environment implementation. In: Computational Collective Intelligence. ICCCI 2017. Lecture Notes in Computer Science, vol. 10449, pp. 325–334. Springer, Nicosia (2017)
12. kubernetes/heapster: compute resource usage analysis and monitoring of container clusters. https://github.com/kubernetes/heapster. Last Accessed 2 July 2018

13. Kubernetes - production-grade container orchestration. https://www.kubernetes.io/. Last Accessed 2 July 2018
14. Lin, S.-W., Ying, K.-C., Lee, C.-Y., Lee, Z.-J.: An intelligent algorithm with feature selection and decision rules applied to anomaly intrusion detection. Appl. Soft Comput. **12**(10), 99–110 (2016)
15. LogicMonitor: SaaS-based performance monitoring platform. https://www.logicmonitor.com/. Last Accessed 1 April 2018
16. Mercl, L., Pavlik, J.: The comparison of container orchestrators. In: Third International Congress on Information and Communication Technology, ICICT 2018, London. Advances in Intelligent Systems and Computing, vol. 797, pp. 677–685. Springer, Singapore (2018)
17. Modi, C., Patel, D., Borisaniya, B., Patel, H., Patel, A., Rajarajan, M.: A survey of intrusion detection techniques in Cloud. J. Netw. Comput. Appl. **36**(1), 42–157 (2013)
18. Nagios - The industry standard In IT infrastructure monitoring. https://nagios.org/, Last Accessed 1 April 2018
19. New Relic: digital performance monitoring and management. https://newrelic.com/. Last Accessed 2 July 2018
20. OpenNMS. https://www.opennms.org/. Last Accessed 2 July 2018
21. Prometheus - monitoring system; time series database. https://prometheus.io/. Last Accessed 2 July 2018
22. PRTG network monitor. https://www.paessler.com/prtg. Last Accessed 2 July 2018
23. Roshan, S., Miche, Y., Akusok, A., Lendasse, A.: Adaptive and online network intrusion detection system using clustering and extreme learning machines. J. Frankl. Inst.-Eng. Appl. Math. **4**, 1752–1779 (2018)
24. Sohal, A.S., Sandhu, R., Sood, S.K., Chang, V.: A cybersecurity framework to identify malicious edge device in fog computing and cloud-of-things environments. Comput. Secur. **74**, 340–354 (2018)
25. Sysdig: kubernetes and docker monitoring, container security and more. https://www.sysdig.com/. Last Accessed 2 July 2018
26. Tsai, P.-H., Hong, H.-J., Cheng, A.-C., Hsu, C.-H.: Distributed analytics in fog computing platforms using tensorflow and kubernetes. In: 19th Asia-Pacific Network Operations and Management Symposium: Managing a World of Things, APNOMS 2017, pp. 145–150. IEEE, Seoul; South Korea (2017)
27. Vieira, K., Schulter, A., Westphall, C.B., Westphall, C.M.: Intrusion detection for grid and cloud computing. IT Prof. **12**(4), 38–43 (2010)
28. Weave Scopu. https://www.weave.works/. Last Accessed 2 July 2018
29. Zabbix - The enterprise-class open source network monitoring solution. https://www.zabbix.com/. Last Accessed 2 July 2018
30. Zhang, Y., Wang, L., Sun, W., Green, R.C., Alam, M.: Distributed intrusion detection system in a multi-layer network architecture of smart grids. IEEE Trans. Smart Grid **2**(4), 796–808 (2011)

Model of the Information System in the Organization for Controlling Current Level of Information Security

Jerzy Stanik, Maciej Kiedrowicz and Jarosław Napiórkowski

Abstract The article outlines a concept of maintaining the required level of security of assets of the information system in the organization (ISO) by making appropriate steering decisions, initiating the generation of the security configurations. The authors proposed and formulated the models of security subject and object as well as the model of the information system in the organization for controlling current level of information security (information recourses) and current performance properties of the operation subsystems, included in the ISO.

Keywords Security subject · Security object · Security environment · Security configurations · Controlling security level

1 Introduction

The multitude ways of perceiving and defining security points to the need of researching it in both scientific and praxeological manner. The purpose of the article is not theoretical deliberation on information security in the organization, but definition of possibilities and method of their use to control the current level of information security in the organization. In the opinion of the authors, the required level of information security may be achieved by making appropriate steering decisions, initiating the generation of relevant security configurations.[1] A possibility of making steering decisions is determined by the following [6, 7]:

[1] Security configuration—a set of security mechanisms of the information processing subsystem or technical or organizational security measures of such subsystem, and relationships between them, reflecting security properties of the security configuration.

J. Stanik · M. Kiedrowicz · J. Napiórkowski (✉)
Faculty of Cybernetics, Military University of Technology in Warsaw, Warsaw, Poland
e-mail: jaroslaw.napiorkowski@wat.edu.pl

J. Stanik
e-mail: jerzy.stanik@wat.edu.pl

M. Kiedrowicz
e-mail: maciej.kiedrowicz@wat.edu.pl

M. Huk et al. (eds.), *Intelligent Information and Database Systems: Recent Developments*, Studies in Computational Intelligence 830,
https://doi.org/10.1007/978-3-030-14132-5_34

– having appropriate models of security subject and object ISO,
– presence of a specific subsystem for controlling performance and security properties distinguished in the subsystems operating in the ISO, e.g. the information processing subsystem, subsystem for using technical and organizational resources, automatic control subsystem, etc.,
– presence of the subsystem in the information processing subsystem for handling security measures and appropriately prepared security mechanisms, e.g. software for protection processes or relevant procedures allowing to generate proper security configurations,
– presence of appropriate procedures of automatic control of its suitability and current level of security.

Figure 1 shows schematic representation of the ISO in terms of controlling its current performance properties and security level. The figure presents three key subsystems. One of them is for controlling security level, second for information processing and third for handling security measures. The following are the elements of the subsystem for controlling security level:

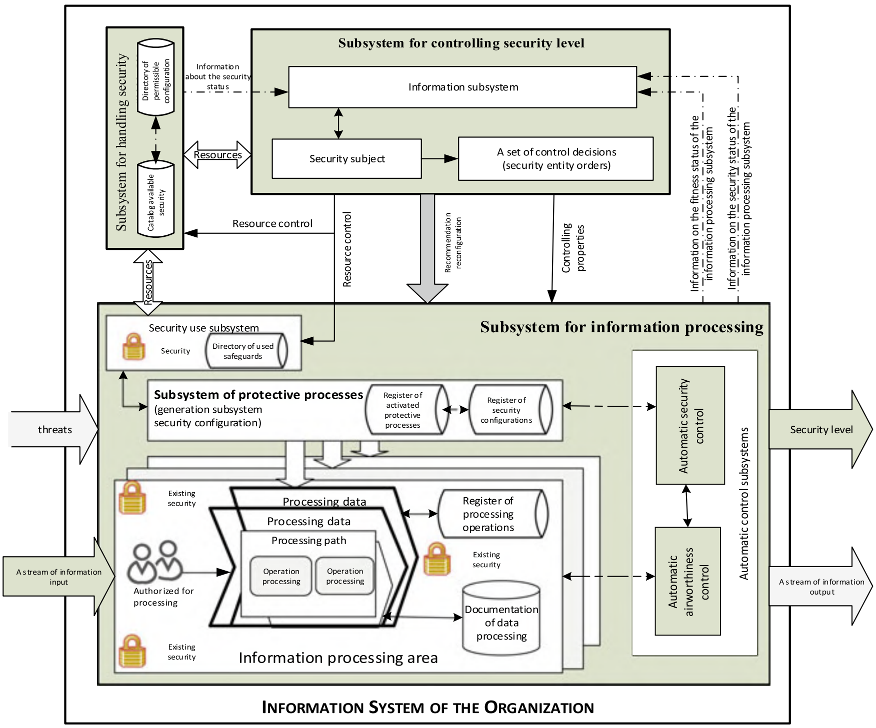


Fig. 1 Illustration of the information system in the organization from the point of view of controlling the information security level

– security subject, also known as the decision-making subject, which consists of a group of officers (security experts) responsible for steering decisions,
– set of numbers (codes) activated by decision-making elements to generate permissible security configuration ensuring the required level of security in the ISO.

When controlling the ISO security level, we are only interested in certain a priori unknown times, when the steering decisions should be made. The ISO states at such times shall be called key states crucial from the point of view of controlling the security level. The occurrence of such a key state is determined by the difference between the value of a current security level and the value of an acceptable security level. In such case, the ISO moves from the safe state to the loss of security. The main cause thereof may be:

1. changes in the ISO security environment, e.g. increased risk stream, new security gaps, higher vulnerability of information resources, etc.,
2. failure and uselessness of resources in particular subsystems included in the information processing subsystem,
3. failure or uselessness of current security mechanisms or security measures of technical or organizational nature in the subsystem for using or handling security measures.

The occurrence of the event specified in points 1, 2, 3 or any combination thereof, further referred to as the emergency situation, is detected by the automatic control subsystems and the information on such event immediately sent to the controlling subsystem. The purpose of the controlling subsystem is to adjust the difference between the current and the required level of security. If the difference is not adjusted by the controlling subsystem, the ISO shall not achieve the specific purpose of operations.

2 Basic Terms and Expressions Concerning the Information System

The following interpretation has been selected from among multiple definitions in the theory of the system engineering as best meeting the requirements of this article: "The system refers to a material object with intentionally organized activity, whose main components, apart from the technical or organizational elements, are also human resources (staff) with predefined operating algorithms" [9].

Following the above definition, the information system in the organization shall be called [8]:

$$SIO =< E, R, DZ > \tag{1}$$

where:

E a finite set of system elements $\{e_j; j \in J\}$, $J = \{1, 2, 3, \ldots, J\}$—a set of indexes in set E,
R a finite set of $\{R_i; i \in I\}$ correlations defined in set $E = \{1, 2, 3, \ldots, I\}$—a set of indexes in set R,
DZ the purpose of system operations.

Set E describing the composition of the information system meets the following criteria:

$$E = \left\{e_j : \xi(j, q), j \in J, q \in Q^j\right\} \tag{2}$$

The $\xi(j, q)$ value shall be construed as the following sentence formula: "The element number $j \in J$ is characterized by feature number $q \in Q^j$, where Q^j refers to a set of indexes of set C^j including features of element number "j"".

The set E including ISO elements may be decomposed in the following manner:

$$E = E^{PS} \cup E^{PPI} \cup E^{PZ} \cup E^{OT} \tag{3}$$

where:

E^{PS} the set of element of the subsystem for controlling security level,
E^{PPI} the set of elements of the subsystem for information processing,
E^{AR} the set of elements of the subsystem for handling security measures,
E^{OT} the set of elements that constitute environment of the subsystem for processing information and controlling security level

In the sets E^{PS} including elements of the subsystem for controlling security level, E^{PZ} of the subsystem for handling security measures and E^{PPI} of the information processing elements, the following functional components may be distinguished:

$$E^{PS} = E^{PS}_{PD} \cup E^{PS}_{PR} \cup E^{PS}_{OT} \tag{4}$$

$$E^{PZ} = E^{PZ}_{PD} \cup E^{PZ}_{PR} \cup E^{PZ}_{OT} \tag{5}$$

$$E^{PPI} = E^{PPI}_{PD} \cup E^{PPI}_{PR} \cup E^{PPI}_{OT} \tag{6}$$

where:

E^{PS}_{PD} the set of elements of the controlling subsystem, which constitute the security subject (decision-making subject),
E^{PS}_{PR} the set of elements of the controlling subsystem, which constitute its subject,
E^{PS}_{OT} the set of elements of the controlling subsystem, which constitute the environment of its subject and object,

E^{PZ}_{PD} the set of elements of the subsystem for handling security measures, which constitute the action subject,
E^{PZ}_{PR} the set of elements of the subsystem for handling security measures, which constitute the subject thereof,
E^{PZ}_{OT} the set of the subsystem for handling security measures, which constitute the environment of its subject and object,
E^{PPI}_{PD} the set of elements of the information processing subsystem, which constitute its subject of processing,
E^{PPI}_{PR} the set of elements of the information processing subsystem, which constitute its subject,
E^{PPI}_{OT} the set of elements of the information processing subsystem, which constitute the environment of its subject and object.

3 Model of the Security Subject of the Information System in the Organization

The security subject is "every intentionally existing and operating subject (individual or collective), analyzed in terms of its security" [4]. Security is most often defined as both the state (achieved security of a given subject) and the process (supporting a sense of security of the subject) [3]. The second approach is more practical, as it reflects more natural and dynamic nature of the concept of security.

The ISO security does not refer to the processing processes related only to the IT infrastructure. It is a much broader term, which also includes the following:

- entirety of processes, during which the information is generated and processed,
- confidentiality, consistency and reliability of actions connected with the collection and processing the data available only to authorized persons,
- the actions aimed at securing the information resources existing on electronic and paper information carriers.

The ISO security subject is a collective entity, being the subject of the processes and events, during which it mentally or physically interacts with potential or already active risks of values that are important therefor. With respect to security analyzed in terms of controlling the current information security level, the security subject, also referred to as the subject of operations, shall be a group of positions—i.e. an element of the automatic process for making steering decisions or the system of automatic security control (SAKB), further referred to as the decision-making process.

The following abbreviations shall be introduced:

PS—a set of ordered pairs: $ps_p = \langle \Theta_p, IT_p \rangle \in \Theta \times 2^{IT}$, hereinafter referred to as the subsystems of operations.

where:

Θ the set of officers determined at the stage of ISO design, included in the information processing subsystem, IT a set of resources, which constitute equipment of the information processing subsystem;

The term s_p state, where p $\in \hat{P}$, for the pth subsystem of operations means the vector of properties describing in detail the current performance or security properties:

$$s_p = \left\langle a_p^q \in A_p^q : p \in \hat{P}, q \in \hat{Q} \right\rangle \tag{11}$$

where:

a_p^q coordinates of the vector of the pth subsystem of operations, describing individual properties,

A_p^q a set of permissible implementation of the qth property of the pth subsystem of operations.

The impact of the subsystems of operations on the state and hence on their performance or security properties may be described in the following manner:

$$\wedge_{\langle p,q\rangle \in \hat{P}\times\hat{Q}}\, a_p^q = a_p^q[u(t)],\, u \in U \tag{12}$$

Therefore, the set of the controllable subsystems of operations may be defined as follows:

$$\widehat{PS} = \left\{ ps_p \in \text{PS} : \vee_{q\in\hat{Q}}[\langle p, q\rangle \in V_u],\, p \in \hat{P} \right\} \tag{13}$$

The purpose of operations of the subsystem for controlling security properties is shown in set $\widehat{PS}$ of controllable subsystems.

In terms of controlling current ISO properties or the information processing subsystem, every subsystem of operations (Fig. 1) may be described in an extended manner:

$$\widehat{ps}_p = \langle np_p, PDZ_p, PRDZ_p, NP_p, CDZ_p \rangle \tag{14}$$

where:

np_p the name of the pth subsystem of operations,

PDZ_p the subject of operations of the pth subsystem of operations,

$PRDZ_p$ the object of operations of the pth subsystem of operations,

NP_p a set of technical and organizational measures that may be invoked or activated in the pth subsystem of operations,

CDZ_p the purpose of operations of the pth subsystem of operations

With the above in mind, the following ordered five shall be used as the ISO security model:

$$\text{PBSIO} = < SPB, PDZPPI, PDZPOZ, ZDS > \tag{15}$$

where:

SPB	a set of security persons, which constitutes the collective ISO security subject,
PDZPPI	the set of persons authorized to process the set of information resources in the ISO,
PDZPOZ	the set of persons authorized to generate a set of permissible security configurations once it is impossible to ensure the required level of security of the information resources in the ISO,
ZDS	a set of steering decisions (operator's recommendations.

Paraphrasing the proposed definition of the security subject, we may assume that the security subject must be aware of the risk and self-aware of participation in such risk.

4 Model of the Security Object of the Information System in the Organization

Following the aforementioned definition of the security subject, both security and its environment are of the object character.[2] In light of the above, the ISO security and security environment become the security object. The ISO security as the object falls neatly into the definition of activities. This approach to ISO security requires defining such terms like: ISO security, security environment, aspects of security and security environment.

The easiest way to define security is to describe it as the state or process, during which a given subject may exist and grow. It is a temporary process or phenomenon. Security also refers to theory and practice aimed at ensuring survival (existence) and exercise of interests of a given subject, in particular through specific objects subject to protection, use of opportunities (favorable circumstances), taking up of challenges, mitigation of risks and prevention (combating) of all types of risks faced by the subject and its interests as well as key processing processes or determined group of protection resources.

The security environment defined as the *security subject* is created by "*external and internal security conditions (conditions allowing the execute the interests of a given subject in terms of security and achievement of the established goals)*" [1]. The aforementioned conditions may be characterized using five basic categories, such as: resources, opportunities, challenges, risks and threats.

The opportunities refer to certain circumstances (phenomena and processes in the ISO security environment) allowing to execute the interests of the security subject and achieve goals. The opportunities are mainly created by the security environment

[2] At this point, we shall not go deeper into the meaning of the word "object", but assume the most general dictionary definition. Therefore, we treat the "object" as the "object of recognition and activities of the subject—i.e. a human".

subjects. On the other hand, the challenges are difficult situations in terms of security, which cause certain dilemmas that need to be confronted by the security subject. An example of a challenge is securing the established group of information resources. The risks refer to uncertainties associated with the subject's own actions and the effects thereof; these are potential negative effects of the subject's own activities on the given subject in terms of security. The threats mean direct and indirect destructive impact on the protected resources or subject of operations. These are the most classic examples of the security environment factors.

According to the information-processing framework, the subject of operations as part of the ISO is the $\widehat{PS}$ set of the subsystems of operations, and the subject of operations is the set of such $e_j \in E^O$ elements, where the purpose of operations of the information processing subsystem is defined. The following may be the elements of the E^{SIO} set like: areas of processing, information processing processes, protection processes, bites of information (information resources) collected or processed as part of the ISO or subsystem of operations with different purposes of operations (e.g.: subsystem for using security measures, subsystems of automatic or automated control).

The key elements of the E^O set are bites of information (information resources) processed by the information processing resources. The aforesaid bites of information are characterized by actual information objects in the organization (e.g. data of natural persons, descriptions of basic assets in the organization, description of supporting assets, etc.). They are formed by the information processing subsystem according to certain rules and are subject to further processing. Every bite of information, hereinafter referred to as the information object (information resource) is marked with number $\in Z^O$ and described using a set of C_z^O feature names.

The following abbreviations shall be introduced:

D the set of steering decisions, hereinafter referred to as the orders, which the security subject (members of the security team) may use to have impact on both current properties of the subsystems of operations, included in the information processing subsystem, and information resources (also including their security attributes);

$\underline{a}(t)$ the vector of the distinguished information resources, whose coordinates define the security states of particular objects (information resources) at time t, e.g. the loss or continuity of confidentiality, integrity, non-repudiation, availability and accountability.

The $a^z(t)$ state of the $\in Z^O$ zth information resource refers to the vector of features describing in detail its current state of security, quality or reliability, etc.:

$$a^z(t) = \left\langle a_q^z(t) \in \ddot{A}_z^q : z \in Z^Z, q \in Q^Z \right\rangle \tag{16}$$

where:

$a_q^z(t)$ coordinates of the vector of the zth state of the information resource, describing individual properties,

$\ddot{A}_z^q$ a set of permissible implementation of the qth property of the pth resource.
Q^Z a set of numbers of the distinguished properties of the resources.

On the basis of the above considerations, it is apparent that the ISO security subject may be either a set or collection of the following elements (objects):

- area of processing,
- information processing process,
- processing operation,
- information resources collected or processed as part of the ISO,
- protection process,
- subsystem of operations with specific purpose of operations, e.g.: subsystem for using security measures, subsystems of automatic or automated control risk analysis subsystem, etc.

5 ISO Model to Control Current Performance or Security Properties

The following ordered five was adopted as the information system model to control current performance properties or current security level of the information resources:

$$\langle SPB, NSA, RKB, OFR, SZQ\rangle \tag{23}$$

where:

SPB the set of positions of officers, appointed within the framework of the organization security services,
NSA the set of various types of emergency situations, which may be signaled by the subsystem of automatic security control,
RKB the family of the permissible security configurations,
OFR the general reconfiguration function,
SZQ the detailed reconfiguration function.

Failure shall be understood as an event that occurred at time t_i, caused by the difference between the desirable security property of the ISO and its current security property. It is in line with the following condition:

$$\wedge\langle p, b\rangle \in \dot{P}xB\colon \dot{\Lambda}\left(w_p^b\right)(t_i) \supset \Lambda\left(w_p^b\right)(t_i) \tag{24}$$

where:

$\dot{\Lambda}\left(w_p^b\right)(t_i)$ the set of the desirable security states of the pth object of the bth type at time t_i
$\Lambda\left(w_p^b\right)(t_i)$ the set of the current security states of the pth object of the bth type at time t_i

To make it possible to compensate for the emergency situation, it is required to determine sets *NSA* and *RKB* at the stage of the ISO design. In case of an emergency situation, the security subject shall make the steering decision to move the ISO from state of the lost security to the security state, while considering the effects of the failure. It means that if another emergency situation occurs, the ISO shall move from the state of the lost security to the security state once again.

After the occurrence of an emergency situation—loss of the required level of security, it is essential to generate permissible or optimum security configuration to efficiently continue the process of safe information processing in the information system of the organization (ISO). The optimal security configuration is generated from among the set of permissible solutions, on the basis of a detailed reconfiguration function of the information security management system, which due to its essence constitutes the criterial function.

Basis on this approach we built model and then real implementation of secret registry model that would provide legal required protection for classified documents stored in secret registry protecting by RFID technology (Fig. 2).

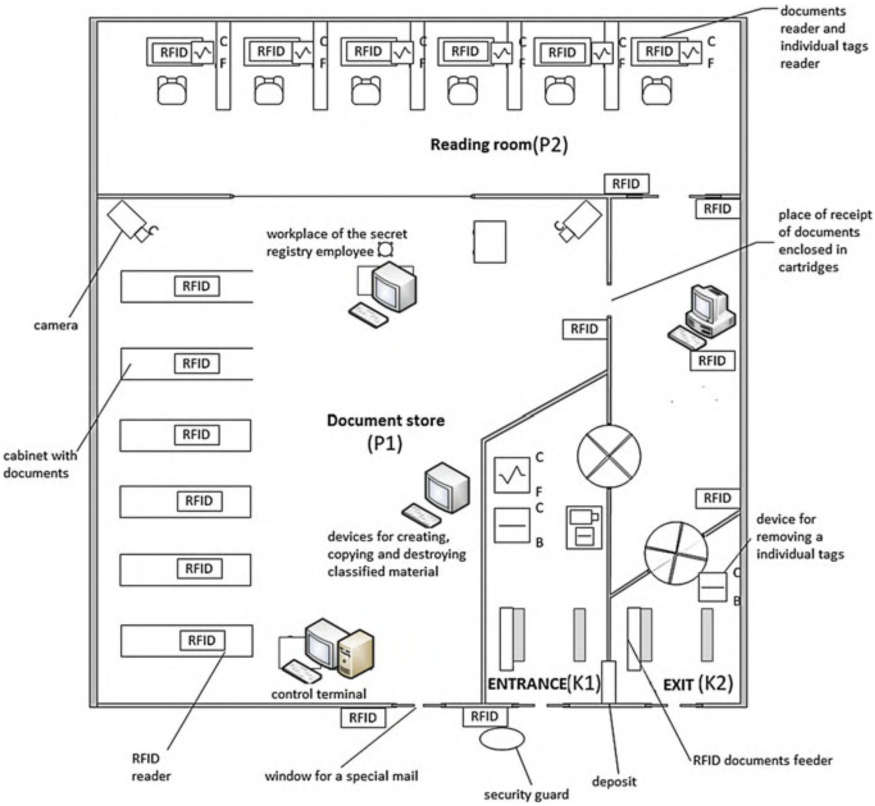


Fig. 2 Diagram of the secret office together with the reading room [2]

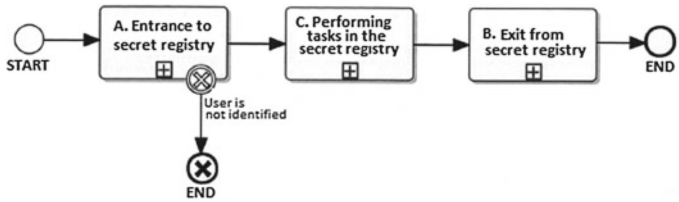


Fig. 3 The general process of the work of secret registry: customer service [2]

In addition to specifying roles and system elements, description of business processes realized in secret office has been developed [2]. For example, as a general process is indicated customer service (Fig 3).

The basic functionalities of our document flow management system include:

- remote identification of unclassified and classified data storage devices labelled for real-time radio-reading at the point of storage and work;
- automatic inventorying of unclassified and classified documents arranged in piles and filed, including automatic detection of relocation;
- control of the movement of data storage devices, as well as unclassified and classified documents across security zones including the unclassified and classified document access control;
- protection against repeated copying of an unclassified and classified document;
- protection of data storage devices and documents against unauthorized relocation.

6 Conclusion

The ISO security and efficiency of the information processing in the organization to a large extent depend on the current qualitative properties of its components (subsystems). In light of the above, it is important to appropriately control the current properties of such subsystems by generating the most desirable performance or security configurations from the set of the permissible solutions, after the emergency situation occurs. The most desirable security configuration is the configuration that not only ensures maintenance of the required level of security, but also provides the best performance properties. This approach allows to use the same for the purpose of maintaining the required level of security of the ISO, through appropriate control on an ongoing basis.

On the basis of the considerations included in this article, the following conclusions may be drawn:

1. To eliminate the effects of failures on the functioning of the ISO, it is well-founded to distinguish three stages:

   a. Definition of the set of permissible security configuration for typical emergency situations,
   b. Development of security configurations in a given emergency situation,
   c. Implementation of the reconfiguration process.

2. The proposed method for controlling current performance or security properties, distinguished as part of the ISO, subsystems of operations, should constitute an integral part of the information management system in the organization.
3. The proposed models of security subject and object may be also used at the stage of design of the information security management system as the "privacy by design" principle recommended according to the GDPR [5].
4. The approach to the issue of security, aimed at the reconfiguration process, results, among other things, from the observations and long-term experience of the authors gained:

   a. during observations of the construction and implementation of such security systems in the organizations and corporations,
   b. during research and implementation projects or research and development activities and seminar discussions concerning information security.

5. Using the results outlined herein, it is recommended to pursue further research in the following areas:

   a. improvement of the structure of security models, including, among other things, intentions and recommendations of the GDPR,
   b. increase of precision of the proposed models by taking a broader perspective of the security environment.

The authors are of the opinion that continuing research in the above-described directions may contribute to better construction of the ISO, with better performance and security parameters.

References

1. Czupryński, A., Wiśniewski, B., Zboina, J.: Bezpieczeństwo. Security. Theory – Research – Practice. CNBOP-PIB, Józefów (2015) (Polish title: Bezpieczeństwo. Teoria – Badania – Praktyka)
2. Napiórkowski, J., Waszkowski, R.: Designing RFID-based systems for government. In: ICIM'2017 The 13th International Conference on Information Management (2017)
3. Piwowarski, J.: Security Studies. Elementary Issues, Scientific Monograph. University of Public and Individual Security "Apeiron" in Cracow, Cracow (2016) (Polish title: Nauki o bezpieczeństwie. Zagadnienia elementarne, Monografia naukowa)
4. Pokruszyński, W.: Security Theory and Practice (academic textbook). Alcide De Gasperi University of Euroregional Economy in Józefów, Józefów (2012) (Polish title: Bezpieczeństwo teoria i praktyka)

5. Regulation (EU) No 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) http://www.giodo.gov.pl/pl/1520284/9745. last accessed 2016/10/30
6. Stanik, J.: Maintaining the required level of security of the integration platform. In: Evaluation of the Usability of the Security Configuration. Newsletter of the Military University of Technology in Warsaw, LXII, No. 1, pp. 181–204, Warsaw (2013)
7. Stanik, J., Kiedrowicz M.: A security service model for maintaining the required level of information security in the organization. In: Simulation in Research and Development, 7(1–2), pp. 35–53, Warsaw (2018) (Polish title: Model służby bezpieczeństwa dla potrzeb utrzymywania wymaganego poziomu bezpieczeństwa informacji w organizacji)
8. Stanik, J., Kiedrowicz, M.: Method for assessing efficiency of the information security management system. In: CSCC 2018 22nd International Conference on Circuits, Systems, Communications and Computers (2018)
9. Stanik, J.: Maintenance of the required level of current functional reliability of the computer automatic control system. Dissertation, Warsaw (1987)

Author Index

M. Huk et al. (eds.), *Intelligent Information and Database Systems: Recent Developments*, Studies in Computational Intelligence 830,
https://doi.org/10.1007/978-3-030-14132-5