INTERNATIONAL CONFERENCE ON

APPLICATIONS OF ARTIFICIAL INTELLIGENCE IN TRANSPORTATION

CO-ORGANIZERS

HO CHI MINH CITY DEPARTMENT OF TRANSPORT

UNIVERSITY OF TRANSPORT & COMMUNICATIONS (UTC)

FENG CHIA UNIVERSITY (FCU)

ASIAN INSTITUTE OF TECHNOLOGY IN VIETNAM

SPONSORS

HO CHI MINH CITY DEPARTMENT OF SCIENCE & TECHNOLOGY

SUMITOMO CONSTRUCTION MACHINERY Co., Ltd

INNOVATIVE TECHNOLOGY DEVELOPMENT CORPORATION

ISBN: 978-604-76-2102-6

International Conference on APPLICATIONS OF ARTIFICIAL INTELLIGENCE IN TRANSPORTATION

SUMITOMO

iTD CORPORATION

AIT-VN

HO CHI MINH CITY - VIETNAM
11 DECEMBER 2020

ISBN: 978-604-76-2102-6

CO-ORGANIZERS

SPONSORS

SUMITOMO

iTD CORPORATION

AIT-VN
Asian Institute of Technology in Vietnam

CONFERENCE PROCEEDINGS

INTERNATIONAL CONFERENCE ON

APPLICATIONS OF ARTIFICIAL INTELLIGENCE IN TRANSPORTATION

NHÀ XUẤT BẢN GIAO THÔNG VẬN TẢI

CONFERENCE PROCEEDINGS

INTERNATIONAL CONFERENCE ON

APPLICATIONS OF ARTIFICIAL INTELLIGENCE IN TRANSPORTATION

Ho Chi Minh City, Vietnam

COMMITTEE

Organizing Committee

1. Assoc.Prof. Nguyen Van Hung
2. Dr. Vo Truong Son
3. MSc. Tran Quang Lam
4. Prof. Tien-Yin (Jimmy) Chou
5. Dr. Phung Van Dong
6. Dr. Ngo Chau Phuong
7. Assoc.Prof. Nguyen Duy Tien
8. Dr. Trinh Quang Khai
9. Dr. Tran Xuan Truong
10. Assoc.Prof. Nguyen Canh Minh
11. Dr. Nguyen Quoc Tuan
12. Dr. Nguyen Xuan Tung
13. Assoc.Prof. Le Van Bach
14. Dr. Nguyen Van Binh
15. Dr. Nguyen Thac Quang
16. MSc. Tran Phong Nha

Program Committee

1. Assoc.Prof. Nguyen Thanh Hai
2. Prof. Le Hung Lan
3. Dr. Bui Ngoc Dung
4. Dr. Nguyen Van Binh
5. Dr. Vu Viet Hung
6. Dr. Dao Thanh Toan
7. Dr. Hoang Van Thong
8. Assoc.Prof. Nguyen Huu Hung
9. Dr. Le Ba Anh
10. Assoc.Prof. Tran Ngoc Hien
11. Dr. Dinh Thi Thanh Binh
12. MSc. Ho Viet Long
13. MSc. Tran Phong Nha
14. MSc. Doan Van Tan
15. Prof. Cheong Siew Ann
16. Prof. Wen-Nung Lie
17. Prof. Tien-Yin (Jimmy) Chou
18. Prof. Dimitar Borisov
19. Prof. Hansung Lee

Secretariat

1. Dr. Vu Viet Hung
2. MSc. Pham Anh Linh
3. MA. Ngo Thuy Linh
4. MSc. Nguyen Thi Mai Nhung
5. MSc. Doan Minh Huy
6. Assoc.Prof. Hoang Thanh Van

Table of Contents

PREFACE

In the recent years, the application of modern technologies to the Transport sector has made great progresses in Vietnam. Infrastructure has been getting smarter. Artificial intelligence has gradually replaced people in the tasks of construction quality assessment, traffic control and management. Compared with humans, artificial intelligence has the advantage of processing speed and capacity, and high objectiveness as well. Therefore, it is able to meet the requirements and execution time of the vehicles' monitoring and operating process and the infrastructure management process.

The University of Transport and Communications is one of the leading institutions in the country in the research on the applications of artificial intelligence in the Transport industry. Many of the university's products have been applied in practice. To strengthen the links in scientific research with domestic and foreign institutions, the University of Transport and Communications has coordinated with Ho Chi Minh City Department of Science and Technology, Ho Chi Minh City Department of Transport, Feng Chia University (Taiwan) and Asian Institute of Technology in Vietnam to organize the International conference "Applications of artificial intelligence in transportation" at the University of Transport and Communications - Campus in Ho Chi Minh City.

The conference has received the attention and contributions of many scientists in the fields of image processing, big data processing, traffic management, logistics and sensor technology.

We hope that the scientific papers and presentations in the conference have brought practical benefits in the development of the Transport industry in general and the artificial intelligence field in particular. At the same time, the conference has also brought a close relationship between scientists in this new field, creating a driving force for development in the coming future.

AITC2020 Organizing Committee

DEVELOPMENT OF AN INTELLECTUAL MONITORING SYSTEM FOR COMPREHENSIVE ASSESSMENT OF URBAN ROAD - TRANSPORT ENVIRONMENTAL POLLUTIONS

Anna Makarova[1*], Andrey Glushko[2], Vitaliy Chelnokov[1], Alexey Matasov[1], Georgii Priorov[2]

[1]*Mendeleev University of Chemical Technology, 9 Miusskaya sq., Moscow, Russian Federation.*

[2]*Institute of Chemical Reagents and Highly Pure Chemical Substances, National Research Centre «Kurchatov Institute», 1 Akademika Kurchatova pl., Moscow, Russian.*

[*]*Corresponding author: 7963819@inbox.ru*

Abstract: This paper systematically analyses a contemporary situation of the environmental impact of waste from year-round maintenance of roads. During the study, a model was created for monitoring and evaluating the effects of anti-icing and wastes that were generated as a result of the maintenance of objects of the road infrastructure, classifies significant typical factors of impact in the developed hierarchical structure for monitoring. The automated system determines the main cause-effect relationships for the subsequent assessment of environmental, socio-economic consequences taking into account the weight of each factor and the parametric dependence of the subcriteria for each factor for a comparative assessment of alternative options. It introduces a specific methodology for a systematic analysis of the socio-economic consequences of the use of agents for maintaining roads for the safety of road transport infrastructure and the safe functioning of society. To implement the developed plan in accordance with this methodology, an analysis of the impact on the main components of the environment, such as: air, water, soil and green spaces, was carried out. The analysis and description of the properties of chemicals which hazardous to the environment and human health and are part of the agents for road maintenance are carried out. A number of research works showed an assessment that was made of the content of chemicals included in the composition of agents for the maintenance of roads in environmental objects, taking into account their migration and transformation in the environment. As a result of study, the paper provides a new method for conducting environmental computer monitoring of the state of environmental objects.

Keywords: information intellectual system, multifactorial ecological-socio-economic assessment, road maintenance waste, environmental monitoring, technogenic-bio-physicochemical impact.

1. INTRODUCTION

The design and implementation of an integrated infrastructure able to combine environmental, social and sensor-driven data in order to enable a real-time management control of the city itself represents an impressive Issue [1]. The development of a modern

methodology and advanced research results in the development and management of intellectual-informational and technological tools for assessing the impact of waste year-round maintenance of roads on the basis of environmental monitoring are associated with the analysis of both technogenic and environmental-socio-economic consequences when maintaining economic facilities [2]. Among the many significant scientific works for solving the tasks set in the project, according to the authors, the following can be distinguished.

Scientific works [3] presented a study of the consequences of the maintenance of territorial economic facilities based on the dialectical economic and environmental geographical approach. The classification of the best options for the dialectical approach includes environmental protection, inter-regional interdependence, etc. Using the example of the use of anti-icing reagents in the city of Moscow, the scientists analyzed and identified factors for assessing the consequences of ecological and geological nature [4]. This work additionally revealed the indirect effect of the waste of roads that interacted with anti-icing - an increase in detergents in the wastewater of car washes during the use of anti-icing. The remains of distributed waste in an environment with anti-icing reagents are investigated taking into account the presence of a number of elements that are harmful to humans and animals and are considered inseparably with other components of road sediment, so it does not show the place and role of anti-icing materials.

A generalized geoecological approach to the analysis of the consequences on the state of the state of natural-social-production systems was presented in the studies [5]. These studies set out the experience of geoecological analysis and assessment of the state of regional and local natural-social-production systems using the example of economic facilities of the Republic of Mordovia. There is a sufficient number of studies based on this approach, including studies during the state environmental review of the technology for winter maintenance of road facilities in Moscow [6]. For a systematic review of the geochemical effect of waste with anti-icing [7], the researchers introduced the term “urban ecosystem”.

However, in the aforementioned studies don’t provide a comprehensive assessment of the interaction of road anti-icing with the urban ecosystem, anti-icing reagents are regarded as products, that are alien to the urban ecosystem, as a result, all factors related to the quality of the road infrastructure and agents for maintaining roads are not identified, studied and summarized and not carried out environmental impact assessment for them [8].

A number of systematic studies was carried out in [9], they identify key factors that affectin the quality of the road, establish the patterns of their interaction at different levels of the hierarchy. Using the decomposition method, a multifactor model is constructed that reflects the main quality indicators of the road infrastructure in the form of a hierarchical system, which shows the place and role of the use of road chemistry materials, including road anti-icing and road impregnating compositions, the multi-level structure of the analytical was developed and jointly considered for the first time monitoring of chemical deicing reagents for an example of their use on the surface of asphalt concrete pavement in winter and summer. The cluster architecture of the quality indicators of the considered objects of road chemistry is determined and a typical structure of the relationship of each quality indicator with the corresponding methods of analytical control and analytical equipment is developed [10].

A multifactor model is based decomposition method, it reflects the main quality indicators of the road infrastructure in the form of a hierarchical system, which shows the place and the role of the use of road chemistry materials, including road anti-icing and road impregnating compositions. The research both introduces a multi-level structure of the analytical and considers monitoring of chemical deicing reagents for an example of their use on the surface of asphalt concrete pavement in winter and summer for the first time.

The results of studies using the systems approach described in [11] present the experience of analyzing and evaluating the management of socio-economic systems based on a combination of quality economics tools and mathematical modeling methods [12]. On the one hand, it presents the results of assessing the impact of standardization, metrology, quality management on socio-economic processes and systems, on the other hand, it estimates the possibility of using the mathematical apparatus and computer technologies to build agent-oriented models for scenario calculations.

The members of the scientific team under the guidance of Academician RAS Meshalkin. formulated the basic principles of the use of chemicals (natural and synthetic) [13]. They introduced the principles of system analysis, mathematization, construction of algorithms, modeling, information intelligent decision support systems and their comparative assessment for environmental systems taking into account socio-economic factors.

2. ENVIRONMENTAL MONITORING METHOD

This research provides a specific method is for conducting environmental computer monitoring of the state of environmental objects, which allows to assess the influence of natural factors successfully and shows the effect of pollutants on the environmental objects. This type of monitoring allows to take into account the synergistic effect of the accumulation of various pollutants that are toxic to humans and the environment.

The technical result consists in increasing the speed of information processing and the correctness of calculations.

The technical result is achieved through the application of the method of analyzing hierarchies during monitoring, including the determination and monitoring of environmental indicators that are sensitive for assessing the environmental situation, the calculation of environmental indicators, and the calculation of environmental indicators in accordance with the hierarchy analysis method, which includes: determination and fixing the presence of each type of pollution in a single month; calculating their average annual detection and fixing the total number of months of detection of each type of pollution; building a matrix for comparing environmental objects in pairs by comparing the main priority indicators; normalization of the matrix by dividing each element of the column by its sum and calculating the average value and weight for each object; pairwise comparison of each type of pollution by the total number of months in which they were recorded for each environmental object; compiling a matrix for determining the total weight for each pollutant; normalization of the matrix by dividing each element of the column by its sum and calculating the average value and weight for each pollutant for each object; formation of a matrix for determining the total weight indicator for each pollutant for all environmental objects; multiplying the formed matrix by a column with the obtained weight indicators of environmental objects; determination of the total percentage weight of each pollutant as a result of the calculations.

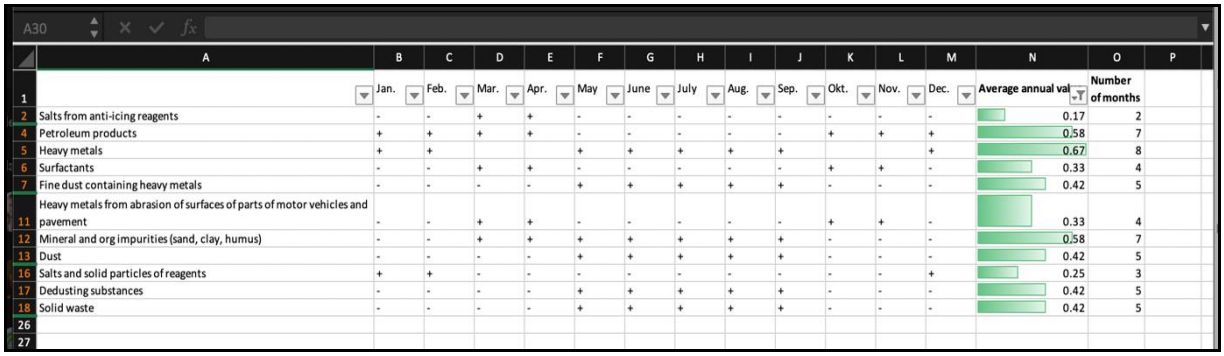
We applied exact methods, including the hierarchy analysis method, to conduct monitoring. This hierarchy analysis method is quite universal for any sphere of human activity: whether it is business, public administration or private life. The analysis of the decision-making problem in the hierarchy analysis method begins with the construction of a hierarchical structure, which includes the goal, criteria and other factors considered that influence the choice.

3. MONITORING ALGORITHM

The monitoring algorithm includes the following stages:

Step 1. Determination and fixing of the presence of each type of pollution in a separate month of the year in various environmental objects: soil and soil, green spaces, water bodies, atmospheric air is carried out.

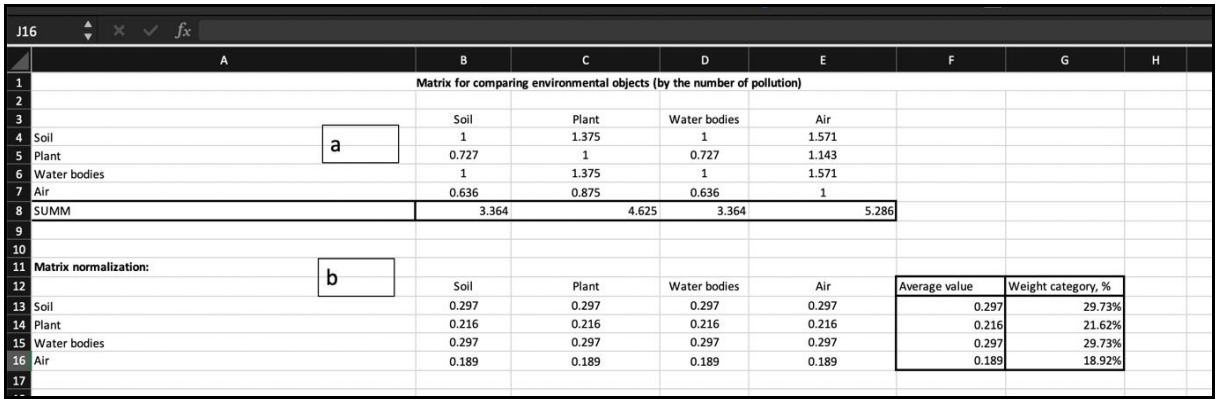
Step 2. Based on the obtained pollution data, an automated extended spreadsheet is compiled for each environmental object, showing the presence of each type of pollution for one month (Figure 1). At the same time, the automated system checks the presence of each type of pollution at the specified environmental object, and, if detected, enters the "+" sign in the cell corresponding to the month of detection. In case of absence or incorrect values, a "-" sign is entered. On the right side of the table, the average annual detection and the total number of months of detection of each type of pollution are calculated.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1		Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sep.	Okt.	Nov.	Dec.	Average annual val	Number of months	
2	Salts from anti-icing reagents	-	-	+	+	-	-	-	-	-	-	-	-	0.17	2	
4	Petroleum products	+	+	+	+	-	-	-	-	-	+	+	+	0.58	7	
5	Heavy metals	+	+			+	+	+	+	+			+	0.67	8	
6	Surfactants	-	-	+	+	-	-	-	-	-	+	+	-	0.33	4	
7	Fine dust containing heavy metals	-	-	-	-	+	+	+	+	+	-	-	-	0.42	5	
11	Heavy metals from abrasion of surfaces of parts of motor vehicles and pavement	-	-	+	+	-	-	-	-	-	+	+	-	0.33	4	
12	Mineral and org impurities (sand, clay, humus)	-	-	+	+	+	+	+	+	+	-	-	-	0.58	7	
13	Dust	-	-	-	-	+	+	+	+	+	-	-	-	0.42	5	
16	Salts and solid particles of reagents	+	+	-	-	-	-	-	-	-	-	-	+	0.25	3	
17	Dedusting substances	-	-	-	-	+	+	+	+	+	-	-	-	0.42	5	
18	Solid waste	-	-	-	-	+	+	+	+	+	-	-	-	0.42	5	
26																
27																

Figure 1. *Automated system element. Monthly contamination table.*

Step 3. As criteria for priority indicators, the total amount of pollution detected in each environmental object during the year is selected. A pair-wise comparison of priority indicators is made and a matrix of comparison of environmental objects by the amount of pollution is compiled Fig. 2, part a. The matrix is normalized by dividing each element of the column by its sum 2, part b. The average value and weight indicators of environmental objects are calculated. The weight of the pollutant shows what percentage of the total amount of pollutants falls on each environmental object.



Matrix for comparing environmental objects (by the number of pollution)

a

	Soil	Plant	Water bodies	Air
Soil	1	1.375	1	1.571
Plant	0.727	1	0.727	1.143
Water bodies	1	1.375	1	1.571
Air	0.636	0.875	0.636	1
SUMM	3.364	4.625	3.364	5.286

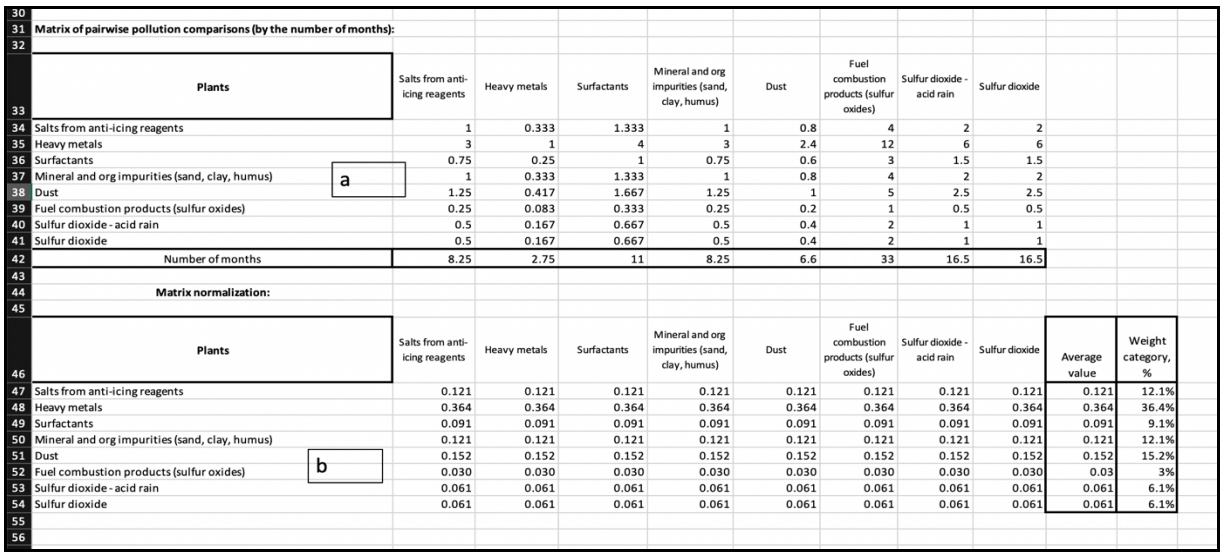
Matrix normalization:

b

	Soil	Plant	Water bodies	Air	Average value	Weight category, %
Soil	0.297	0.297	0.297	0.297	0.297	29.73%
Plant	0.216	0.216	0.216	0.216	0.216	21.62%
Water bodies	0.297	0.297	0.297	0.297	0.297	29.73%
Air	0.189	0.189	0.189	0.189	0.189	18.92%

Figure 2. *Comparison matrix of environmental objects with its normalization, calculated average value and category weight.*

Step 4. Construction of a pollution comparison matrix by pairwise comparison of each type of pollution by the total number of months in which they were recorded for each environmental object individually Fig. 3, part "a". The matrix is normalized by dividing each element of the column by its sum Fig. 3, part "b". The average value and weight for each pollutant for each object is calculated. The weight of the pollutant shows what percentage of the total amount of pollution falls on the environment.



Matrix of pairwise pollution comparisons (by the number of months):

a

Plants	Salts from anti-icing reagents	Heavy metals	Surfactants	Mineral and org impurities (sand, clay, humus)	Dust	Fuel combustion products (sulfur oxides)	Sulfur dioxide - acid rain	Sulfur dioxide
Salts from anti-icing reagents	1	0.333	1.333	1	0.8	4	2	2
Heavy metals	3	1	4	3	2.4	12	6	6
Surfactants	0.75	0.25	1	0.75	0.6	3	1.5	1.5
Mineral and org impurities (sand, clay, humus)	1	0.333	1.333	1	0.8	4	2	2
Dust	1.25	0.417	1.667	1.25	1	5	2.5	2.5
Fuel combustion products (sulfur oxides)	0.25	0.083	0.333	0.25	0.2	1	0.5	0.5
Sulfur dioxide - acid rain	0.5	0.167	0.667	0.5	0.4	2	1	1
Sulfur dioxide	0.5	0.167	0.667	0.5	0.4	2	1	1
Number of months	8.25	2.75	11	8.25	6.6	33	16.5	16.5

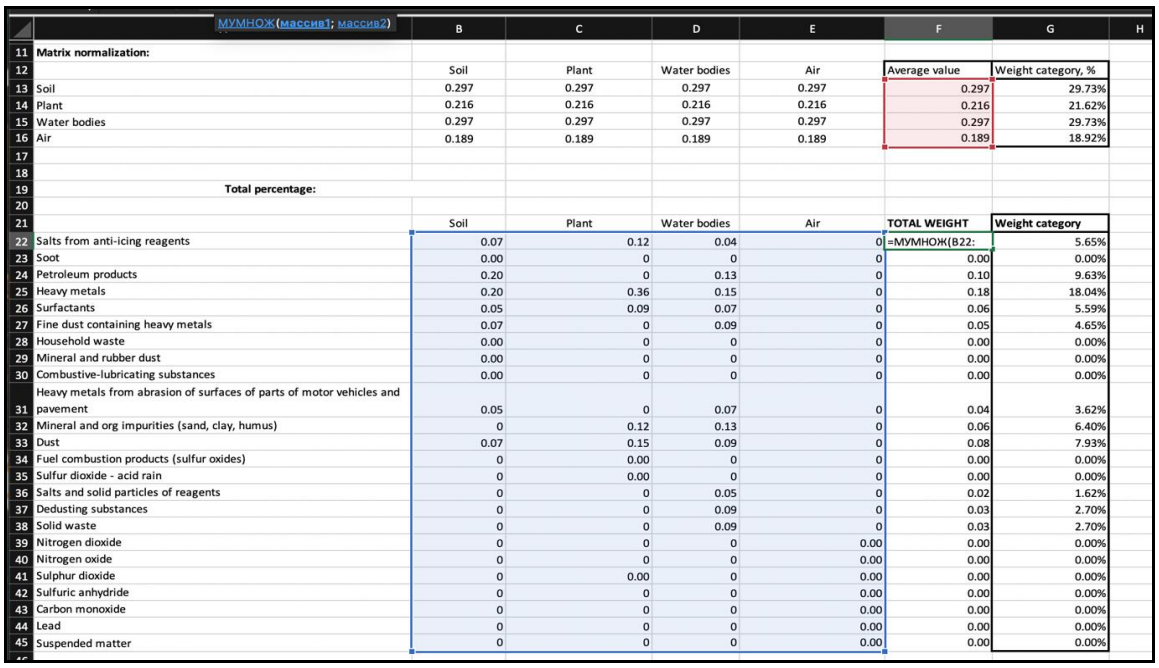
Matrix normalization:

b

Plants	Salts from anti-icing reagents	Heavy metals	Surfactants	Mineral and org impurities (sand, clay, humus)	Dust	Fuel combustion products (sulfur oxides)	Sulfur dioxide - acid rain	Sulfur dioxide	Average value	Weight category, %
Salts from anti-icing reagents	0.121	0.121	0.121	0.121	0.121	0.121	0.121	0.121	0.121	12.1%
Heavy metals	0.364	0.364	0.364	0.364	0.364	0.364	0.364	0.364	0.364	36.4%
Surfactants	0.091	0.091	0.091	0.091	0.091	0.091	0.091	0.091	0.091	9.1%
Mineral and org impurities (sand, clay, humus)	0.121	0.121	0.121	0.121	0.121	0.121	0.121	0.121	0.121	12.1%
Dust	0.152	0.152	0.152	0.152	0.152	0.152	0.152	0.152	0.152	15.2%
Fuel combustion products (sulfur oxides)	0.030	0.030	0.030	0.030	0.030	0.030	0.030	0.030	0.03	3%
Sulfur dioxide - acid rain	0.061	0.061	0.061	0.061	0.061	0.061	0.061	0.061	0.061	6.1%
Sulfur dioxide	0.061	0.061	0.061	0.061	0.061	0.061	0.061	0.061	0.061	6.1%

Figure 3. *Example of a pollution comparison matrix for plants.*

Step 5. Formation of a matrix for determining the total weight indicator for each pollutant for all environmental objects Fig.4. For each cell, the percentage annual weight of each element of pollution in a single environmental object is selected. If the automated system does not find the weight indicator for the cell, the system resets the cell. The total weighting factor of the pollutant is calculated as the average for environmental objects.

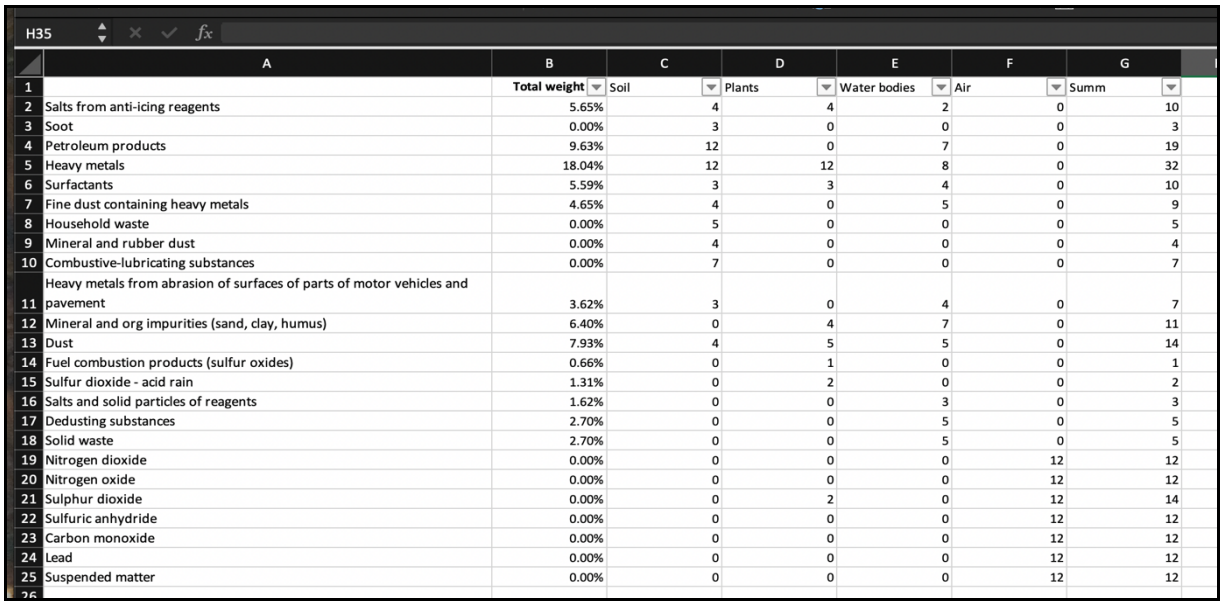


	B	C	D	E	F	G
Matrix normalization:						
	Soil	Plant	Water bodies	Air	Average value	Weight category, %
Soil	0.297	0.297	0.297	0.297	0.297	29.73%
Plant	0.216	0.216	0.216	0.216	0.216	21.62%
Water bodies	0.297	0.297	0.297	0.297	0.297	29.73%
Air	0.189	0.189	0.189	0.189	0.189	18.92%
Total percentage:						
	Soil	Plant	Water bodies	Air	TOTAL WEIGHT	Weight category
Salts from anti-icing reagents	0.07	0.12	0.04	0	=МУМНОЖ(B22:	5.65%
Soot	0.00	0	0	0	0.00	0.00%
Petroleum products	0.20	0	0.13	0	0.10	9.63%
Heavy metals	0.20	0.36	0.15	0	0.18	18.04%
Surfactants	0.05	0.09	0.07	0	0.06	5.59%
Fine dust containing heavy metals	0.07	0	0.09	0	0.05	4.65%
Household waste	0.00	0	0	0	0.00	0.00%
Mineral and rubber dust	0.00	0	0	0	0.00	0.00%
Combustive-lubricating substances	0.00	0	0	0	0.00	0.00%
Heavy metals from abrasion of surfaces of parts of motor vehicles and pavement	0.05	0	0.07	0	0.04	3.62%
Mineral and org impurities (sand, clay, humus)	0	0.12	0.13	0	0.06	6.40%
Dust	0.07	0.15	0.09	0	0.08	7.93%
Fuel combustion products (sulfur oxides)	0	0.00	0	0	0.00	0.00%
Sulfur dioxide - acid rain	0	0.00	0	0	0.00	0.00%
Salts and solid particles of reagents	0	0	0.05	0	0.02	1.62%
Dedusting substances	0	0	0.09	0	0.03	2.70%
Solid waste	0	0	0.09	0	0.03	2.70%
Nitrogen dioxide	0	0	0	0.00	0.00	0.00%
Nitrogen oxide	0	0	0	0.00	0.00	0.00%
Sulphur dioxide	0	0.00	0	0.00	0.00	0.00%
Sulfuric anhydride	0	0	0	0.00	0.00	0.00%
Carbon monoxide	0	0	0	0.00	0.00	0.00%
Lead	0	0	0	0.00	0.00	0.00%
Suspended matter	0	0	0	0.00	0.00	0.00%

Figure 4. *Matrix for determining the total weight indicator.*

Step 6. Multiplying the formed matrix by a column with the obtained weight indicators of environmental objects and determining the total percentage weight of each pollutant as a result of the calculations.

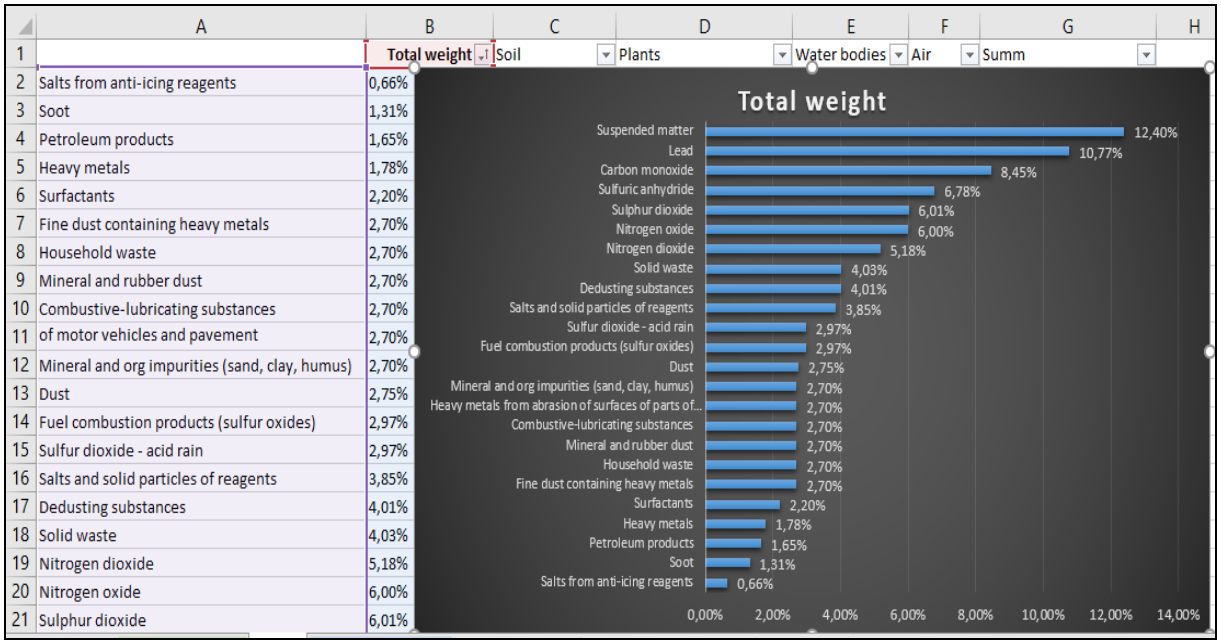
To emphasize the totals, the results of automated analytical calculations were duplicated on the tab "Summary" Fig. 5. For the convenience of monitoring studies, data on the total number of months in which pollution was recorded are added to the matrix of totals.



	A	B	C	D	E	F	G
		Total weight	Soil	Plants	Water bodies	Air	Summ
2	Salts from anti-icing reagents	5.65%	4	4	2	0	10
3	Soot	0.00%	3	0	0	0	3
4	Petroleum products	9.63%	12	0	7	0	19
5	Heavy metals	18.04%	12	12	8	0	32
6	Surfactants	5.59%	3	3	4	0	10
7	Fine dust containing heavy metals	4.65%	4	0	5	0	9
8	Household waste	0.00%	5	0	0	0	5
9	Mineral and rubber dust	0.00%	4	0	0	0	4
10	Combustive-lubricating substances	0.00%	7	0	0	0	7
11	Heavy metals from abrasion of surfaces of parts of motor vehicles and pavement	3.62%	3	0	4	0	7
12	Mineral and org impurities (sand, clay, humus)	6.40%	0	4	7	0	11
13	Dust	7.93%	4	5	5	0	14
14	Fuel combustion products (sulfur oxides)	0.66%	0	1	0	0	1
15	Sulfur dioxide - acid rain	1.31%	0	2	0	0	2
16	Salts and solid particles of reagents	1.62%	0	0	3	0	3
17	Dedusting substances	2.70%	0	0	5	0	5
18	Solid waste	2.70%	0	0	5	0	5
19	Nitrogen dioxide	0.00%	0	0	0	12	12
20	Nitrogen oxide	0.00%	0	0	0	12	12
21	Sulphur dioxide	0.00%	0	2	0	12	14
22	Sulfuric anhydride	0.00%	0	0	0	12	12
23	Carbon monoxide	0.00%	0	0	0	12	12
24	Lead	0.00%	0	0	0	12	12
25	Suspended matter	0.00%	0	0	0	12	12

Figure 5. *Totals Matrix.*

Based on the final matrix, a histogram of all pollutants is compiled Fig. 6, taking into account their percentage prevalence in environmental objects and their environmental load.



	A	B	C	D	E	F	G	H
1		Total weight	Soil	Plants	Water bodies	Air	Summ	
2	Salts from anti-icing reagents	0,66%						
3	Soot	1,31%						
4	Petroleum products	1,65%						
5	Heavy metals	1,78%						
6	Surfactants	2,20%						
7	Fine dust containing heavy metals	2,70%						
8	Household waste	2,70%						
9	Mineral and rubber dust	2,70%						
10	Combustive-lubricating substances	2,70%						
11	of motor vehicles and pavement	2,70%						
12	Mineral and org impurities (sand, clay, humus)	2,70%						
13	Dust	2,75%						
14	Fuel combustion products (sulfur oxides)	2,97%						
15	Sulfur dioxide - acid rain	2,97%						
16	Salts and solid particles of reagents	3,85%						
17	Dedusting substances	4,01%						
18	Solid waste	4,03%						
19	Nitrogen dioxide	5,18%						
20	Nitrogen oxide	6,00%						
21	Sulphur dioxide	6,01%						

Figure 6. *Total weight diagram.*

4. CONCLUSION

The implementation of the claimed method of conducting environmental computer monitoring of the state of the environment can be carried out at each computerized workplace. A research strategy has been worked out, research tasks have been set, a theoretical basis has been determined, a systematic analysis of heuristic-computational tools for formalizing heterogeneous data and knowledge of an intelligent information system for multivariate environmental, socio-economic assessment and monitoring of the effects of year-round waste of roads on household facilities and the environment on all stages of the life cycle, taking into account their engineering and technological effectiveness, direct, indirect, prolonged and delayed effects.

A technique has been developed to analyze the impact of the maintenance of public roads on the natural environment. The dynamics of changes in the car flow, the length of roads and road safety for 1990-2019 is revealed. The analysis of the existing regulatory framework for monitoring environmental impact during the year-round maintenance of roads. The system of placing monitoring points depending on the category of roads, land category, the presence of roadside forest belts with the determination of the zones of environmental impact by various groups of pollutants is justified. As a result of the studies, it was justified the location of observation points depending on the category of road, the conditions of its location, the factors requiring concentration of the number of observation points were identified, the list of controlled pollutants by main environmental components was clarified, the frequency of sampling during operation of roads of different categories was recommended.

Also, the influence of the degree of abrasiveness of solid road sediment de-icing reagents on the acceleration of wear of asphalt concrete pavement, tires and shoes was revealed.

ACKNOWLEDGMENT

The reported study was funded by RFBR according to the research project № 18-29-24185.

REFERENCES

[1] L. Calderoni (ed), Distributed Smart City Services for Urban Ecosystems, Alma Mater Studiorum, Università di Bologna, 2015. https://core.ac.uk/download/pdf/76519845.pdf

[2] I. Allegrini, F. Costabile. An intelligent transport system based on traffic air pollution control. Air Pollution XII, C.A. Brebbia (ed.), WIT Press, 2004. ISBN 1-85312-722-1.

[3] M. Kh. Mamedov. Functional economic and environmental structural development as an object-sustainable regional environment. International Journal of Applied and Basic Research. 2016, 4(2), P. 398-403 [in Russian]

[4] V.A. Korolev, V.N. Sokolov, E.N. Samarin, Assessment of the ecological and geological consequences of the use of deicing reagents in Moscow. Engineering geology. 2009, 1, P. 34–43 [in Russian]

[5] A. A. Yamashkin, A. V. Kiryushin, A. K. Kovalenko et all. Geoecological analysis of the state of natural-social-production systems. A.A. Yamashkin et all (ed.). Saransk. Publishing House of Mordov. University, 2004, 260 p. [in Russian]

[6] Conclusion of the expert commission of the state environmental review of the draft technical documentation “Technology for winter cleaning of the carriageway of highways, streets, driveways and squares (road facilities in Moscow) using anti-icing reagents and crushed granite fraction 2-5 mm (for winter periods from 2010-2011 and beyond), approved by Order of the Department of the Federal Service for Supervision of Natural Resources in the Central Federal District of September 12, 2011 # 08-E. 47p. [in Russian]

[7] A.N. Glushko, A.M. Bessarabov, G.G.Priorov. CALS-system of ecological monitoring of chemical anti-icing materials on the major environmental components. Chemical Engineering Transactions. 2018. 70. P. 451-456. DOI: 10.3303/CET1870088 http://www.aidic.it/cet/18/70/076.pdf

[8] M.A. Zanini, F. Faleschini, J.R. Casas. State-Of-Research on Performance Indicators for Bridge Quality Control and Management. Front. Built Environ., 2019, DOI:10.3389/fbuil.2019.00022

[9] A.N. Glushko, V.V. Chelnokov, A.M. Bessarabov. Computer quality management system for assessing the environmental impact of deicing reagents on the environment. Mathematical methods in engineering and technology - MMTT. 2016. 10 (92). P. 176-177. [in Russian]

[10] G.G. Priorov, A.S. Makarova, V.V. Chelnokov, A.N. Glushko, A.V. Matasov. Analytic hierarchy process of geoecological monitoring of the impact of road infrastructure on environmental objects. 19th International Multidisciplinary Scientific GeoConference SGEM 2019. 2019. 19

(2.1). P. 745-752. DOI: 10.5593/sgem2019/2.1/S08.097.

[11] S. N. Kuzmina, N. V. Androsenko. Using the methods of mathematical modeling and tools of the economy of quality to ensure sustainable development of socio-economic systems. Journal of Science of Science, 2014, 6 (25)

[12] V.B. Korobov, A.G. Tutygin, Advantages and disadvantages of the method of analysis of hierarchies, News of the Herzen State Pedagogical University of Russia, Russia, 2010, 122, P 108-115, [in Russian]

[13] V.P. Meshalkin, Yu.A. Ivashkin, M.A. Nikitina. Computer-aided multiagent model of chemicophysiological processes in the human gastrointestinal tract as a living biochemical system, Doklady Chemistry, Great Britain, 2019, 484 (1), pp 29-32, DOI: 10.1134/S0012500819010087.

CRACK DETECTION FOR ASPHALT PAVEMENTS USING DEEP LEARNING-BASED FASTER R-CNN

Van Phuc Tran[1], Thai Son Tran[1], Van Phuc Le[2], Hyun Jong Lee[1*]
[1]*Department of Civil and Environmental Engineering, Sejong University, South Korea*
[2]*Faculty of Civil Engineering, Campus in Ho Chi Minh City, University of Transport and Communications, No 450-451 Le Van Viet Street, Tang Nhon Phu A Ward, District 9, Ho Chi Minh City, Vietnam.*
[*]*Corresponding Author: hlee@sejong.ac.kr (H.J. Lee)*

Abstract: In pavement management system, evaluating the pavement condition is an essential step in determining the appropriate rehabilitation strategy to use for a pavement. Pavement condition evaluation includes crack identification which serves an important role in assessing the in-situ condition of the pavement based on the existing surface distresses. Currently, crack identification has been done widely by using visual inspection method which is time-consuming. In this study, a deep learning technique is used to detect and classify different types of cracks occurred in asphalt pavements. Crack images captured from scanned road image are classified into three categories composed of fatigue, longitudinal and transverse cracks. A total 20,000 images of 1,852x1,000 pixels are collected to establish a database for training deep learning model. The training images are labelled based on the three categories of crack type and trained using a deep learning-based framework called Faster region-based convolutional neural networks (R-CNN). The trained Faster R-CNN model is validated using scanned pavement surface images obtained by scanning 1,200m road section. The validation shows the detection and classification accuracy of the trained Faster R-CNN model is 89.6%. It can be concluded that Faster R-CNN is a promising technique for crack detection in pavement.

Keywords: pavement distress, automated distress detection, deep learning, faster R-CNN.

1. INTRODUCTION

Pavement condition evaluation is necessary in order to select suitable pavement rehabilitation method [1]. Pavement condition evaluation includes crack identification which serves an important role in assessing the in-situ condition of the pavement based on the existing surface distresses. Currently, crack determination has been done by using visual inspection method which is time-consuming and uneconomical.

To overcome the disadvantages of the visual inspection method, researchers have conducted several studies to identify pavement cracks automatically. In the past decades, image-based algorithms of crack detection have been investigated widely; thresholding [2] edge detection [3] and mathematical morphology [4] are the most popular approaches among the algorithms. Recently, deep learning has been proven as a powerful technique for image classification and object detection and therefore has been applied to detect crack on the pavement. For instance, Mandal et al. [5] used deep convolutional neural networks (DCNN)

to detect road crack automatically. The author utilized images acquired from mobile cameras to make dataset for training and testing deep learning model. Yusof et al. [6] used DCNN for crack detection on asphalt pavement. The authors classified images into four classes including non-crack, longitudinal, transverse, and fatigue crack then these images were used for training and testing the DCNN model. However, DCNN can only be used for image classification which is impossible to differentiate various crack types within one image. Hence, it is necessary to propose another approach to improve the accuracy of crack detection and classification crack type.

In this study, Faster R-CNN [7] which is developed for object detection was applied to build a detection tool identifying crack using surface images of asphalt pavement. Crack images obtained from a digital camera were classified into three categories including fatigue, longitudinal, and transverse cracking. Furthermore, twenty thousand images which is labelled based on the crack type with a size of 1,852x1,000 pixels were used to establish a training database. The database was then trained using the Faster R-CNN and validated using 2,400 images of 1,852x1,000 pixels resolution for crack detection.

2. BACKGROUND

In 2015, Ren et al. [7] developed a deep learning framework named Faster R-CNN, which became one of the famous object detection architecture. Faster RCNN was developed based on previous deep learning architectures including CNN [8], R-CNN [9], and Fast R-CNN [10] Faster R-CNN is composed of three main parts, which are convolution layers, region proposal network (RPN), and classes, bounding boxes prediction. Therein, convolution layers are used to extract features from an original input image or the feature maps in a deep CNN [11]. RPN is a small neural network sliding on the last feature map of the convolution layers and predict whether there is an object or not and predict the bounding box of those objects [7]. Finally, fully connected neural networks takes as an input the regions proposed by the RPN and predict object class (classification) and Bounding boxes (Regression) [7]. Figure 1 shows the architecture of Faster R-CNN.

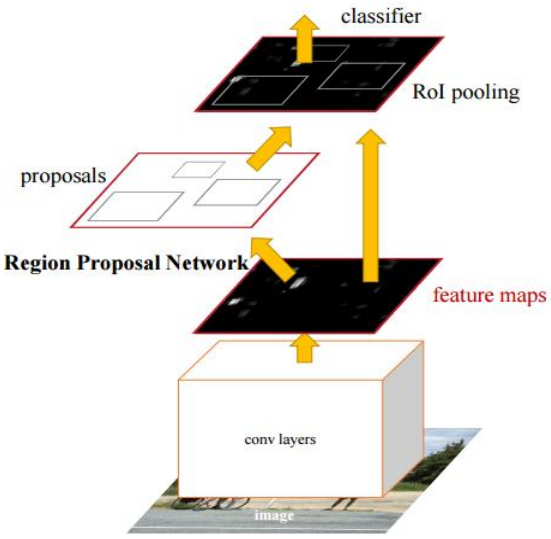


Figure 1. *Architecture of Faster RCNN [7].*

3. METHODOLOGY

The process conducted in this study is shown in Figure 2. In summary, surface images from the field survey are divided to small patches with size of 1,852x1,000 pixels. These smaller images are then labelled with bounding boxes with their corresponding crack type. Then, these labelled images were used to train the Faster R-CNN model. During training, the best trained network model is obtained and used further in the analysis.

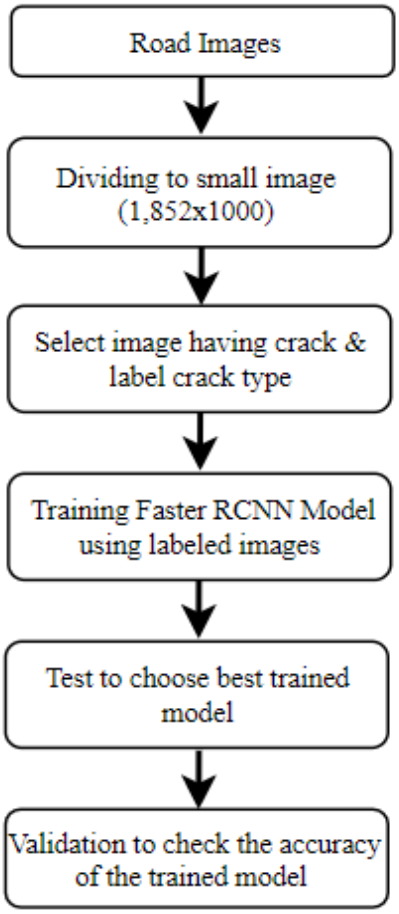


Figure 2. *The entire process for automating detection program.*

4. TRAINING AND TESTING

4.1. Training

To train the Faster R-CNN network, a total of 20,000 images with a size of 1,852x1,000 pixels were used in this study. Figure 3 shows the typical images used for training. The numbers of training images for each object are summarized as follows:

Fatigue cracking:	4,762 objects
Longitudinal cracking:	4,825 objects
Transverse cracking:	3,036 objects

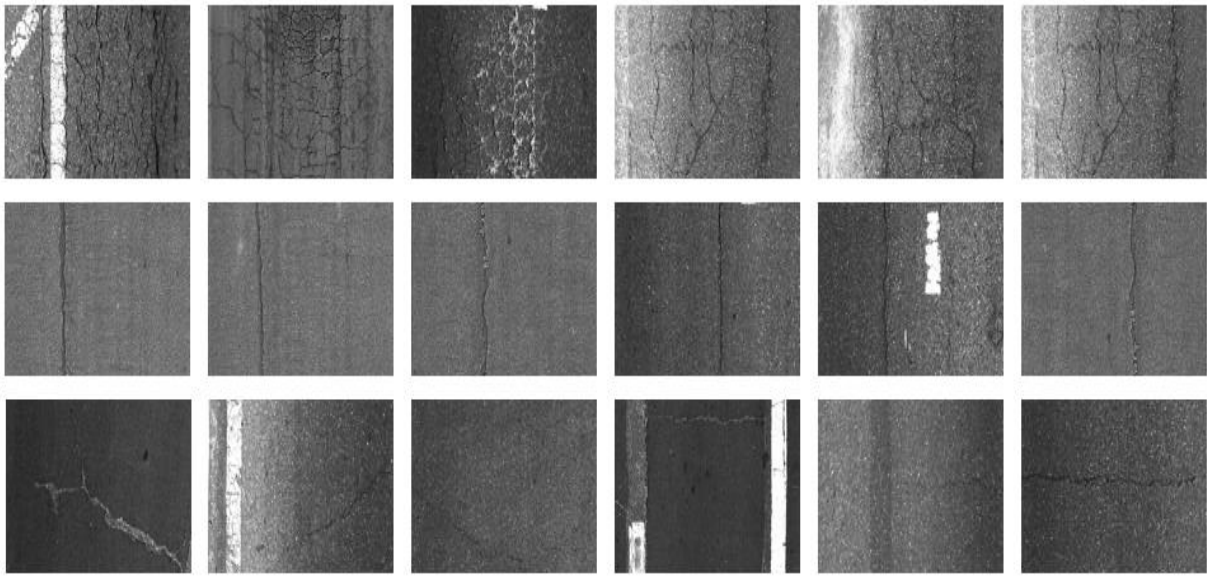


Figure 3. *Typical training patterns of crack distress.*

4.2. Testing

After one circle of training, a trained model will be generated automatically. Usually, the training process is repeated to obtain a good model (set of parameters) for prediction therefore, several models were generated during training. Each model was checked to determine which shows the best performance. In detail, 1,000 small images which are not used in training were used in testing to select the best model.

Figure 4 shows a typical result after using the best trained network for crack detection. It can be seen from the figure that each detected distress is bounded by a rectangular box with a specific colour representing each crack type (i.e., yellow, green, and red colours respectively represent longitudinal, transverse and fatigue cracks).

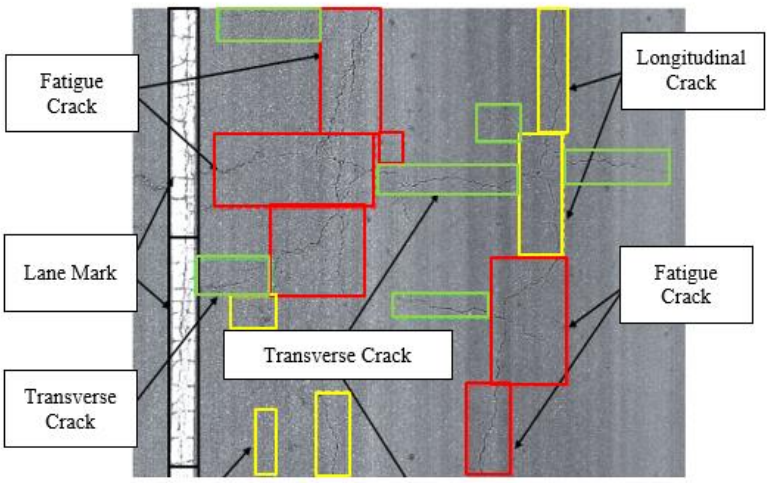


Figure 4. *Typical crack detection results obtained from the trained network.*

5. VALIDATION OF THE BEST TRAINED MODEL

A 1,200m road section (120 original images with size of 3,704x1,000 pixels) was surveyed to validate the performance of the trained model. First, the images were detected to identify crack using the trained model automatically. Then, manual checking was applied to check the accuracy of the trained model. Table 1 shows the validation result of detection of the proposed method. The detection considering crack type have accuracy 89.6% by using the trained model from Faster RCNN architecture.

Table 1. *Prediction Accuracy of the Network Model Considering Crack Type Only.*

Type of Cracking	**Longitudinal (m)**	**Transverse (m)**	**Fatigue (m^2)**
Quantity	1094.47	558.56	488.59
Error (%)	10.4	9.8	11.1
Accuracy (%)	89.6	90.2	88.9
Average Accuracy (%)		89.6	

6. CONCLUSION

In this study, from road scanned image, a deep learning framework based Faster RCNN was used to detect crack considering its three main types including fatigue, longitudinal, and transverse crack occurred on the asphalt pavement. More than 20,000 images containing various crack patterns were used for training Faster RCNN network. The best model selected after training was tested and validated using pavement surface images obtained from 1,200m of a road section. It was observed from the validation study that the network model proposed in this study performed with accuracy 89.6%. It can be concluded that Faster RCNN is a promising approach to detect distress on the pavement.

During training and testing, it was found that the amount of data for training have a significant effect on the performance of the trained model. Generally, the trained model can detect with higher accuracy as the number of training data increasing.

In the future, it is necessary to detect crack considering both its type and severity level as well as detect other distresses on the pavement.

REFERENCES

[1] M. S Zafar, S. N. Shah, M. J. Memon, T. A. Rind, M. A. Soomro, Condition Survey for Evaluation of Pavement Condition Index of a Highway, Civil Engineering Journal, 5(2019), 1367-1383. 10.28991/cej-2019-03091338

[2] H. Oliveira, P. L. Correia. Automatic road crack segmentation using entropy and image

dynamic thresholding, 2009 17th European Signal Processing Conference, 2009, 622-626.

[3] H. Zhao, G.Qin, X. Wang, Improvement of canny algorithm based on pavement edge detection. ed. 2010 3rd International Congress on Image and Signal Processing, 2010, 964-967.

[4] N. Tanaka, K. Uematsu, A Crack Detection Method in Road Surface Images Using Morphology. MVA, 1998, 17-19.

[5] V. Mandal, L. Uong, Y. Adu-Gyamfi, Automated road crack detection using deep convolutional neural networks. ed. 2018 IEEE International Conference on Big Data (Big Data), 2018, 5212-5215.

[6] N. A. M. Yusof, A. Ibrahim, M. H. M. Noor, N. M. Tahir, N. M. Yusof, N. Z. Abidin, M. K., Deep convolution neural network for crack detection on asphalt pavement. ed. Journal of Physics: Conference Series, 2019, 01, 2020.

[7] S. Ren, Faster r-cnn: Towards real-time object detection with region proposal networks. ed. Advances in neural information processing systems, 2015, 91-99.

[8] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks. ed. Advances in neural information processing systems, 2012, 1097-1105.

[9] R. Girshick, Rich feature hierarchies for accurate object detection and semantic segmentation. ed. Proceedings of the IEEE conference on computer vision and pattern recognition, 2014, 580-587.

[10] R. Osman Girshick, Fast r-cnn. ed. Proceedings of the IEEE international conference on computer vision, 2015, 1440-1448.

[11] H. Zhang, Image prediction for limited-angle tomography via deep learning with convolutional neural network. arXiv preprint arXiv:1607.08707.

APPLICATIONS OF ARTIFICIAL INTELLIGENCE IN MANUFACTURING SYSTEMS

Tran Ngoc Hien[1*]

[1]*University of Transport and Communications, No 3 Cau Giay Street, Hanoi, Vietnam.*

[*]*Corresponding author: tranhien.tkm@utc.edu.vn*

Abstract: Artificial intelligence (AI) has been applied efficiently in many fields such as smart manufacturing, smart finance, smart education, smart healthcare, smart city, autonomous vehicles, etc. This paper presents an overview of applications in manufacturing field in which focuses on the model of smart manufacturing systems with applying AI, core technologies for implementation as well as benefits when apply AI for carrying out smart manufacturing. AI based smart manufacturing concepts including cognitive manufacturing, reconfigurable manufacturing, and autonomous manufacturing enable the manufacturing systems adapt to the changes of manufacturing environment autonomously. This ability allows reducing the downtime as well as suit with the customised products. For implementation, core technologies including cognitive technology, agent technology, swarm intelligence and information and communication technology (ICT) infrastructure have been applied.

Keywords: smart manufacturing, artificial intelligence, swarm intelligence, cognition, ant colony, agent technology.

1. INTRODUCTION

Artificial intelligence (AI) based systems enable to have intelligent behaviours like human behaviours such as decision making, reasoning, learning and cooperating with other entities [1,2]. Currently, AI has been applied in many areas from education, healthcare, economy, and digital government to industrial fields such as robotics, manufacturing, and autonomous equipments. The AI based systems show more intelligent and efficiency than the traditional systems.

In the manufacturing field, Industry 4.0 is characterized by the autonomous systems with cyber and physical representation and advance in ICT such as cloud computing, internet of things, internet of service, agent technology, cognitive technology, big data, data mining, and artificial intelligence. Smart manufacturing systems (SMS) built with intelligent cyber-physical systems (CPSs) are the heart of the fourth industrial revolution [3]. SMS aim at improving the levels of intelligence and autonomy due to the CPSs for adapting to the manufacturing changes. Industry 4.0 is expected to generate new business models, as well as increase the efficiency of manufacturing processes in which CPS based SMS can react flexibly to changes. SMSs have autonomic self-characteristics such as self-configuration, self-monitoring, and self-healing.

Currently, manufacturing demands are changing from mass production to mass customization, which is focused on exclusive and individualized demand. This trend requires the manufacturing systems to have the ability to adapt to the fast change of the manufacturing environment. To face this challenge, the new trend of manufacturing system development is to apply autonomous behaviors inspired from living systems and nature to have the sustainable manufacturing systems. Industry 4.0 enables the realization of sustainable manufacturing using the ubiquitous ICT infrastructure. The development of manufacturing not only considers technological innovations but also economic, social and environmental effects. In the literature, many novel paradigms known as intelligent manufacturing systems (IMS) have been proposed such as holonic [4], biological [5], reconfigurable [6], and cloud manufacturing systems [7].

In traditional manufacturing systems, the workers play an important role, with their knowledge and experience, the workers adapt flexibly to changes in the manufacturing environment. With new problems, their knowledge is updated through learning. Information processing model of human is shown in Figure 1.

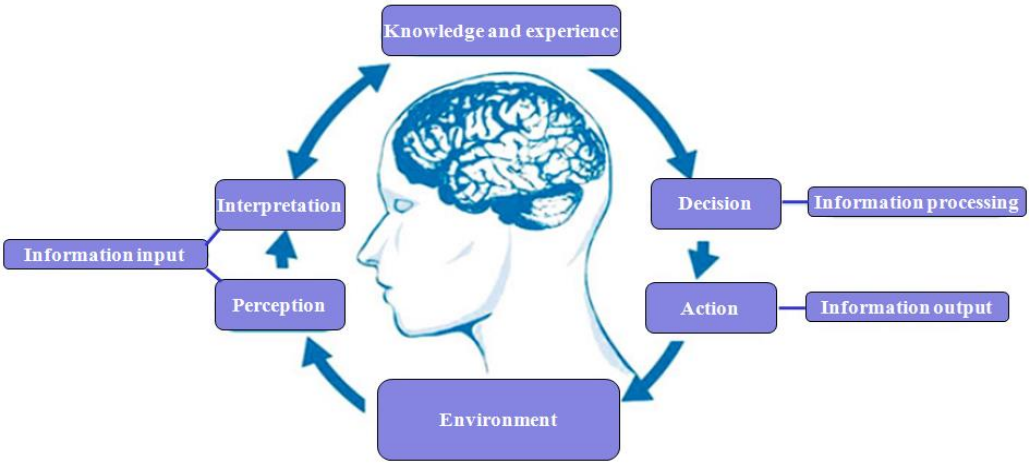


Figure 1. *Information process model of human [8].*

With abilities in processing information and cognitive abilities, the workers keep an important role in monitoring, control, production planning. The workers with the ability to solve problems and cognitive capacity enable to adapt with changing manufacturing environment. However, the system manipulated by workers with a high price is only suitable for small production.

In the era of computer integrated manufacturing, workers are replaced by automatic control systems and robots, so the cognitive abilities of workers in solving problems such as perception, learning, and reasoning for a decision also are removed. The limit of the automatic control system is not capable to adapt to the changes due to the system operating under preset programs. So the system needs to reset and restart when an error occurs. To overcome these shortcomings, and combination of both advanced automatic control system and the cognitive abilities of human, cognitive sciences and artificial intelligence as well as the biology inspired technologies have been applied to the manufacturing systems which make the production system to become more intelligent and more flexible.

The paper presents the application of biology inspired technologies allowing the system to perform intelligent behaviours which mimic human activities such as reasoning for decision-making, cooperation in problem solving. Figure 2 shows an example of the manufacturing system with applying the achievements of cognitive science. In this system, the machines work together for machining workpiece. The system adapts flexibly with the change of orders and changes in the machining shop.

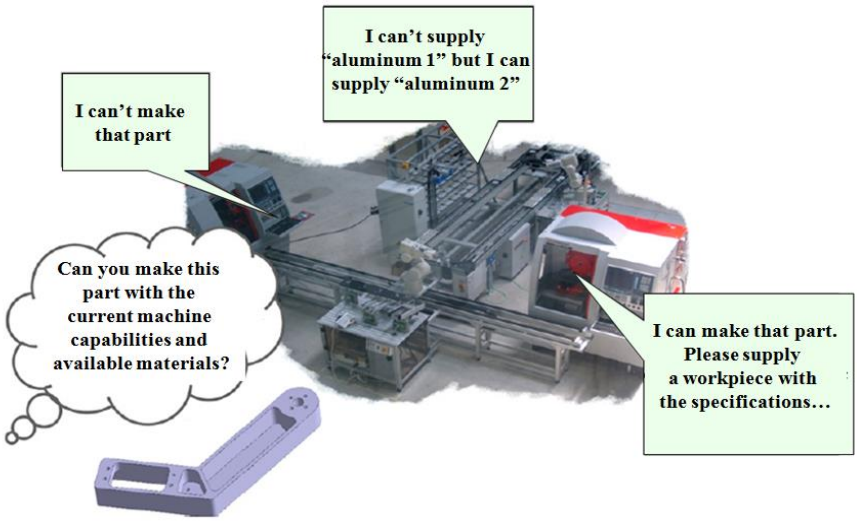


Figure 2. *A machining system with applying the cognitive science [9].*

2. CORE TECHNOLOGIES

2.1. Cognitive technology

In order to overcome the shortcomings and to combine the advantages of both, automated systems and cognitive capabilities of human, cognitive technology is applied into manufacturing systems. In which cognitive architecture for implementing the control system is used [10].

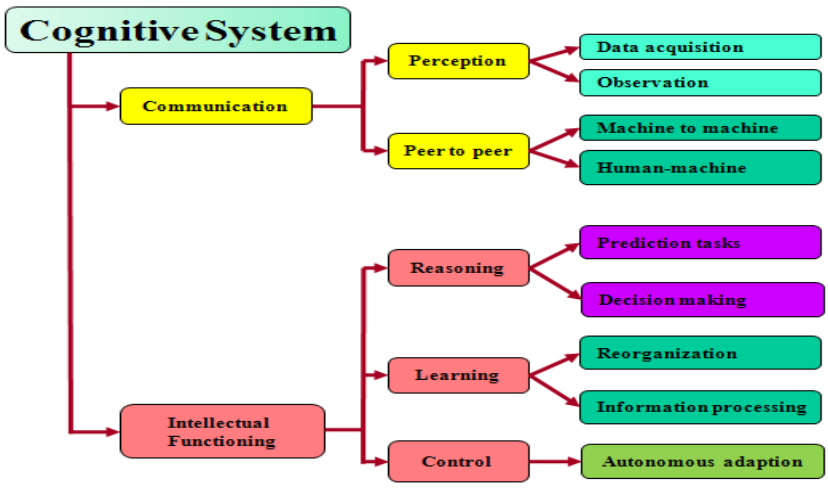


Figure 3. *Cognitive system.*

The cognitive system was proposed by Zhao & Son (2008) [10]. In this system, the BDI architecture inspired from cognitive science was applied. This architecture is based on a human decision-making model from cognitive science that comprises knowledge models, methods for perception and control, methods for planning, and a cognitive perception-action loop.

The paradigm "cognition" in terms of the manufacturing system denotes that machines and processes are equipped with cognitive capabilities and cognitive controls in order to enable them to increase their scope of autonomous operations. A cognitive control consists of three general actions: perceiving information in the environment, reasoning about those perceptions using existing knowledge, and acting to make a reasoned change to the environment. Cognitive capabilities such as perception, reasoning, learning and planning turn the technical system into ones that "know what they are doing" [11]. Manufacturing systems with cognitive capabilities will be much easier to interact and cooperate with, and they will be more robust, flexible and efficient.

In cognitive manufacturing systems, two aspects of cognition are considered to apply into manufacturing environments that are communication and intellectual functions as shown in Figure 3. The first aspect considers methods for data acquisition, observation; connecting machine to machine, machine and human. The second aspect is intellectual functions such as learning, reasoning, and control.

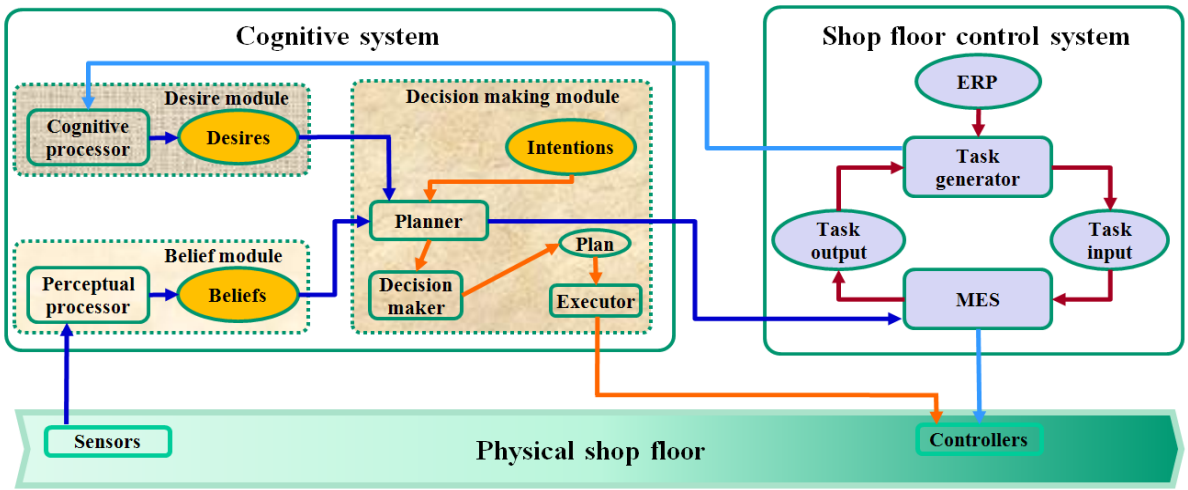


Figure 4. *Cognitive system for controlling a shop floor [10].*

Figure 4 illustrates the cognitive system for controlling the shop floor. This framework is based on the BDI architecture of human decision making. In this framework, the cognitive system is responsible for monitoring and controlling the shop floor. The task generator receives orders from the enterprise resource planning (ERP) and then generates sequences of tasks, which are required to produce the orders. These tasks are sent to both of the manufacturing execution system (MES) and the cognitive system. The belief module is responsible for updating information on the shop floor. These data are then compared with the desires in the desire module by the decision making module. In case the data match with the desired goals, a message is sent to the MES to report the normal status and the shop floor continues running. Otherwise the data are different from the desired goals because of an error

happening. The decision module generates a new plan and controls the shop floor. The MES controls the shop floor when the machines or processes are recovered.

2.2. Swarm intelligence

In the nature, swarm intelligence is built from simple interactions of individuals and with their environment. Swarm intelligence is expressed through mechanisms to adapt to environmental change which shows in the study of biology, such as ant colony, fish school, birds, and the cooperation of human in problem solving.

Ant colony shows swarm intelligence as finding the shortest path from food source to their nests through simple interactions among them using a chemical substance called "pheromone" as shown in Figure 5 [12]. Ant colony adapts to environmental changes by changing the relationship between the members. Applying this principle to the manufacturing field in which the manufacturing system is considered as a combination of autonomous and collaborative entities, the manufacturing system adapts with disturbances by comparing the ability of the machine to the requirements of the product. The disturbances in manufacturing systems are the unplanned changes such as the tool wear, the machine break down, and so on. Each machine has a value of "pheromone" to pass a particular type of disturbances, and the machine having the highest pheromone value is selected to perform the work of the breakdown machine.

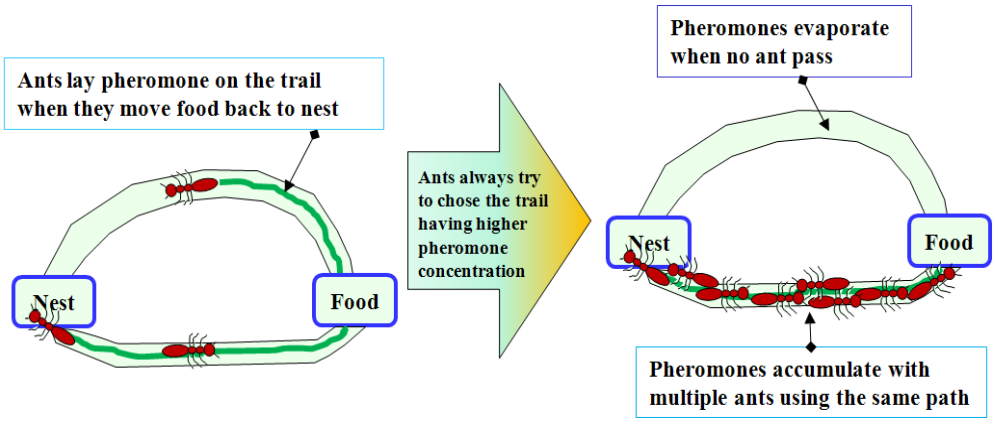


Figure 5. *Ant colony mechanism for finding the shortest path.*

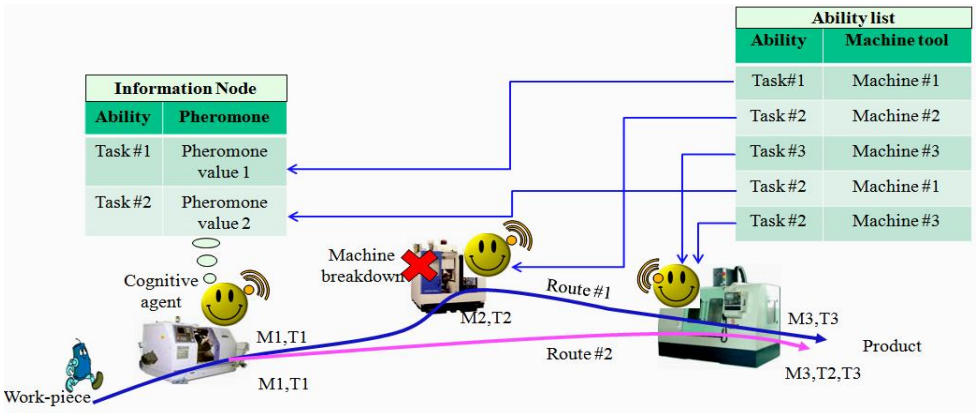


Figure 6. *Disturbance adaptation of a machining system.*

Figure 6 illustrates the cooperation among the cognitive agents in the machining shop. The machining system includes three machining centers. Machining sequence, the initial value of the processing time and the pheromone values of the tasks are assumed. Assuming the task#2 can be performed on any machine. At the start, the system sends tasks to the corresponding machine based on the machine code. The machining system activates the route # 1 in which the machine #1 (M1), machine #2 (M2), and machine #3 (M3) carry out the task #1 (T1), task #2 (T2), and task #3 (T3). In case the machine #2 is broken, route #2 is activated to keep the system running.

3. AI BASED MODELS OF SMART MANUFACTURING SYSTEMS

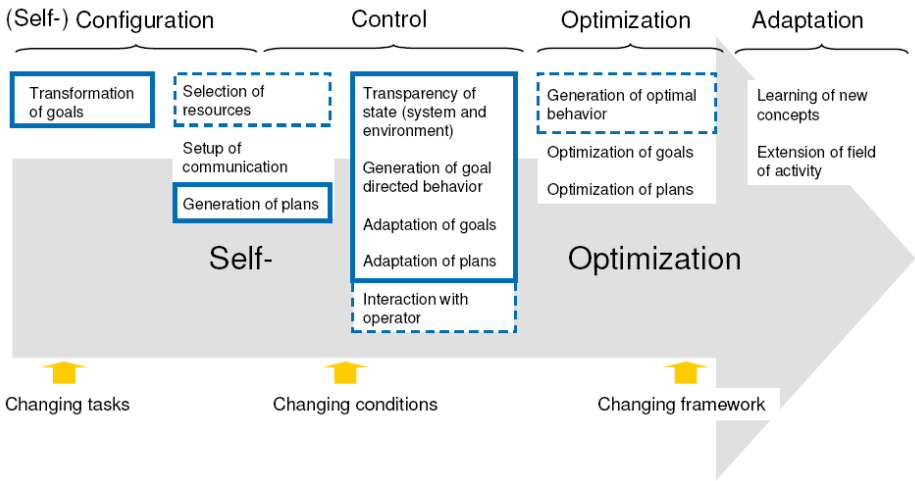


Figure 7. *Aspects of self-X characteristics at the cell level [13].*

The modern manufacturing systems show the self-X capability such as self-configuration, self-control capability, self-optimization, and self-adaptation. Figure 7 describes aspects of self-X characteristics at the cell level. These capabilities can be classified into reconfiguration, autonomy, and intelligence of the manufacturing systems. According to the requirements to having the advanced characteristics of manufacturing systems, new concepts not only for the machine level but also for the system level have been proposed in the literature.

The modern manufacturing systems require the reconfigurable characteristic. Reconfiguration is to rearrange and restructure manufacturing resources that require the rescheduling method and reconfigurable ability of manufacturing systems. A dynamic rescheduling is done when there is an occurrence of disturbances such as the machine breakdown, malfunction of robot or transporter with long recovering time. Here, a new schedule is generated when the current schedule is affected by disturbances. Reconfigurable manufacturing system (RMS) is designed for rapidly adjusting its production capacity and functionality by rearranging or changing its components to adapt to the changes of the manufacturing environment. Unlike the flexible manufacturing system (FMS), the RMS does not have a fixed capacity and functionality, and it is designed through the use of reconfigurable hardware and software. The RMS could be reconfigured both on the overall system's structure level and on the machine level.

The new trend in manufacturing filed is to apply the bio-inspired technologies to equip the machines and processes with autonomous behaviours. The new concepts in manufacturing have been proposed such as genetic, biological, holonic, and intelligent manufacturing system with biological principles. Autonomy allows the system to recover autonomously without either upper level aids such as the ERP, and the MES or the operator intervention. The autonomous method is suitable for the changes that are not necessary to reschedule. In this method, each entity in the manufacturing system is an autonomous entity so that it can overcome the disturbances by itself or communicate with the others to overcome the disturbances. Reactive and collaborative methods were proposed following this criterion. Reactive method is an autonomous control of an entity to overcome disturbances by itself, while the collaborative method is used for a cooperation of an entity with other entities in order to adapt to disturbances. In order to implement reactive/collaborative methods, the distributed control architecture is required. The control architecture changes from centralized control of non-intelligent entities in hierarchical structures of the FMSs towards decentralized control of intelligent entities in distributed structures.

In the real manufacturing system, the information systems such as ERP, and MES keep the main role for operating the manufacturing system. The ERP generates the process planning for manufacturing the product. The MES is responsible for generating the schedule to control and monitor the shop floor. At normal status, the shop floor is controlled by the existing system as MES which dispatches the tasks to the corresponding machines and monitors the status of the shop floor. In case the disturbance happens, the agent system keeps the shop floor running using the generated solution for overcoming the disturbance by the self-adjustment or cooperation mechanism of agents. The connection of the agent system to the existing information system is shown in Figure 8. The agent system percepts the status of the shop floor and makes a decision in the case of the disturbance. If the agent system can not handle the disturbance, the MES generates a new schedule and controls the shop floor.

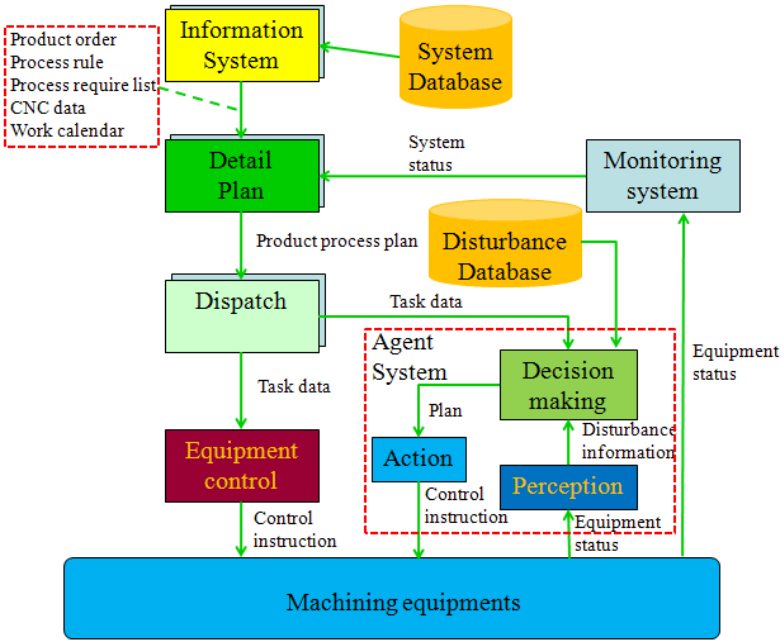


Figure 8. *The integration of the agent system with the existing systems.*

4. CONCLUSIONS

Currently, the manufacturing system should be stopped for repairing the damaged machines or processes caused by the disturbance. The limitations of this method are the reduction of productivity, the decrease in utilization of machining shop, and the effect of approach depends on the experience of operators. To overcome the shortcomings of this method, AI based systems enable to adapt to disturbances. The AI based systems increase the productivity of the manufacturing system due to the reduction of the downtime of the system. This method keeps the manufacturing system running in the case of disturbance. The objectives of the AI based system for adapting to disturbances are as follows:

- Allowing the control system to take an action during disturbances and continue to operate instead of stopping the manufacturing system completely.
- Equipping the entities in the manufacturing system with the decision making and self-controlling abilities.

REFERENCES

[1]. National artificial intelligence strategy, Smart Nation Digital Government Office, Singapore, 2019.
[2]. Global artificial intelligence industry whitepaper, China, 2019.
[3]. W. Wahlster, Industrie4.0: Cyber-physical production systems for mass customization. German Czech Workshop Ind. 2016, 4, 55.
[4]. P. Leitao, A bio-inspired solution for manufacturing control systems. In International Conference on Information Technology for Balanced Automation Systems; Azevedo, A., Ed.; Springer: Boston, MA, USA, 2008; pp.303–314.
[5]. K. Ueda, I. Hatono, N. Fujii, and J. Vaario, Reinforcement learning approaches to biological manufacturing systems. Ann. CIRP 2000, 49, 343–346.
[6]. P. Leitao, J. Barbosa, and D. Trentesaux, Bio-inspired multi-agent systems for reconfigurable manufacturing systems. Eng. Appl. Artif. Intell. 2012, 25, 934–944.
[7]. O. Fisher, N. Watson, L. Porcu, D. Bacon, M. Rigley, and R. L. Gomes, Cloud manufacturing as a sustainable process manufacturing. J. Manuf. Syst. 2018, 47, 53–68.
[8]. Christian Brecher - Integrative production technology for high-wage countries, Springer-Verlag Berlin Heidelberg, 2012.
[9]. S. Kristina, E. Christoph, G. Thomas, and A. Farhad, Design-to-fabrication automation for the cognitive machine shop, Advanced Engineering Informatics 24, 2010, 251–268.
[10]. X. Zhao and Y. Son, BDI-based human decision-making model in automated manufacturing systems, Int J Model Simul, 28(3), 2008, 347-356.
[11]. M. F. Zaeh, M. Beetz, K. Shea, et al., The cognitive factory, In: EIMaraghy, H.A (ed) Changeable and reconfigurable manufacturing systems, Springer, 2009, 355-371.
[12]. N. H. Tran, Development of an autonomous manufacturing system for adapting to disturbances, PhD Dissertation, Ulsan, Korea, 2011.
[13]. K. Tobias, H. Werner, and B. Christian, SOAR-based sequence control for a flexible assembly cell, Proceeding ETFA'09 Proceedings of the 14th IEEE international conference on Emerging technologies & factory automation, 2009.

DEVELOPMENT OF A SMART APPLICATION FOR SAFE INLAND WATERWAY TRANSPORTATION BASED ON BUILT-IN SENSORS AND GPS

Duc-Nghia Tran[1], Ton That Long[2], Nguyen Canh Minh[3], Pham Tien Thanh[4], Duc-Tan Tran[5*]

[1]*Institute of Information Technology, Vietnam Academy of Science and Technology, Vietnam*

[2]*International University, Vietnam National University, Hochiminh City, Vietnam*

[3]*University of Transport and Communications, No 3 Cau Giay Street, Hanoi, Vietnam.*

[4]*VNU University of Engineering and Technology, Hanoi, Vietnam*

[5]*Faculty of Electrical and Electronic Engineering, Phenikaa University, Hanoi, Vietnam*

[*]*Corresponding author: tan.tranduc@phenikaa-uni.edu.vn*

Abstract: In this paper, we present an Android application, namely BoatSafe which offers easy and effective navigation of waterways in Vietnam. The data were collected directly from both GPS and the accelerometer sensor, then these data were processed to determine the user's "Sailing on Water" status through the motions and the position. The software offers some important features such as: provide the positions, velocities, and heading angle of vehicles at the present time; save the route for a certain time period; identify the outage of GPS signal; collision warning feature, identify a collision and call a pre-selected phone number in the event of a collision.

Keywords: inland waterway transport (IWT), global positioning system (GPS), android.

1. INTRODUCTION

Vietnam has about 2,360 rivers and canals with a total length of 220,000 km. Throughout the network, 19% of the total length (about 41,800 km) are accessible, and 7% of it (≈15,436 km) are managed and operated. The total length of main roads in the North and the South is about 4,553 km. The State of Vietnam manages 65 waterways in the North, 21 in the Central and 101 in the South. Inland waterway transport plays an important role in the operation of Vietnam's economy, which is endowed with two large deltas [1]. Most of the distance from the inland waterway in Vietnam is less than 200 km. Each loading method has the appropriate distance to compete with acceptable operating costs and gain more market share from other types of loading. Inland Waterway Transport (IWT) in Vietnam dominates shipments with a distance of less than 300 km and especially with a distance of 100-200 km. Within this range, IWT accounted for a controlling market share of 81% in tons.

Actual water transport in Vietnam is mixed, overlapping traffic between maritime, inland waterways, many types of vehicles circulating on the river, especially in the areas of Hai Phong port, Quang Ninh, and Cuu Long river delta. While domestic and foreign ships are always fully equipped with support systems as prescribed by the Vietnam Maritime Code and

the International Maritime Organization (IMO), etc…, most types of Inland waterway facilities of Vietnam are not yet fully equipped. Especially, the small type of experimental vehicles mainly operates the ship by the traditional method, based on the experience of the operator. If weather factors are taken into account, for IWT vehicles operating along the coast, the route to the island and on the routes connecting the islands will still have a high risk of accidents without support systems.

Currently, to locate ships/boats, the device is mainly satellite (GPS is the most popular) [2-5]. The error of GPS positioning is mainly due to the following 6 reasons (excluding artificial error SA has been turned off): 1) Ephemeris data; 2) Satellite clock; 3) Delay in the ionosphere; 4) Delay in the troposphere; 5) Multi-path noise; 6) Receiver (including software). It can be clearly seen that in bad weather conditions, the GPS signal is often weak and even the loss of GPS signal causes the navigation process to be interrupted. Normally, to support GPS, one of the solutions is to use the inertial navigation system (INS). The INS has two outstanding advantages when compared to other navigation systems: 1) its ability to operate autonomously and 2) high accuracy in short periods of time. The combination of GPS and INS is best because the INS will support GPS very effectively [6-7].

Smartphones have become very popular in recent years. Initially, smartphones included the features of conventional mobile phones combined with other popular devices such as GPS, magnetic compass sensor, inertial sensor (used to build built INS system from this sensor) already built into the machine. In Vietnam, the monitoring of waterways is mostly based on commercial GPS devices [2-4]. Position monitoring of IWT boat/vessel is essential for the management and administration. Large boats/vessels nowadays have to have electric or magnetic compasses, or both, to show the direction of the train. For phone development, there are Google Map (for all applications - not just inland waterways) and a number of other applications (primarily using GPS and magnetic sensors). In this paper, we proposed to use smartphone as a device that integrates navigation, journey monitoring, velocity determination, driving directions for inland watercraft, serving the management of these boats/vessels.

2. MATERIAL AND METHODS

2.1. Data collection and detection of the status “Sailing on Water”

Methods recently proposed for classification in animals/people/vehicles are mainly based on different machine learning algorithms such as decision-trees, k-means, HMMs, and SVM. In this paper, we develop a decision-tree (DT) that uses built-in sensors from the smartphone. Data are collected directly from both GPS and the accelerometer sensor, then processed to determine the user's “Sailing on Water” status through the motions and the position, as described by the flow chart in Figure 1).

Firstly, we use the Elevation service which provides elevation information for locations on the surface of the earth, including depth locations on the ocean floor. In the cases that Google cannot possess exact elevation measurements at the precise location which is requested, the Elevation service will interpolate and provide an average value using the four nearest locations. The Elevation Service object communicates with the Google Maps API Elevation Service, which receives elevation requests and returns elevation data. As shown in

Figure 1, we can determine if the user is in water or land. In case that the user is in water (i.e. on boat, ship, vessel, etc.), the acceleration data are needed.

The raw data stream from the accelerometers is the acceleration of each axis X, Y, Z in the units of g-force [8-9]. Most of the existing accelerometers provide a user interface to configure the sampling frequency so that the user could choose the most suitable sampling rate through experiments. After the raw data have been collected from the accelerometer sensor, the next step is to pre-process these data before performing any further statistical computations. One purpose of the data pre-processing is to reduce the noise from the sensors. The absolute acceleration is a good candidate to discriminate between behaviors with high dynamic and low dynamic movements:

$$A_d[i] = (AX^2[i] + AY^2[i] + AZ^2[i])^{0.5}, \quad (1)$$

where i is time index, AX, AY, AZ are the accelerations in X, Y, and Z directions respectively. If value of A_d is greater than a threshold (denoted by Threshold_A_d), the status of the user is assigned as "Sailing on Water". Consequently, a series of tasks will be executed, such as provide the positions, velocities, and heading angle of vehicles at the present time; save the route for a certain time period; identify the outage of GPS signal; collision warning feature, identify a collision and call a pre-selected phone number in the event of a collision.

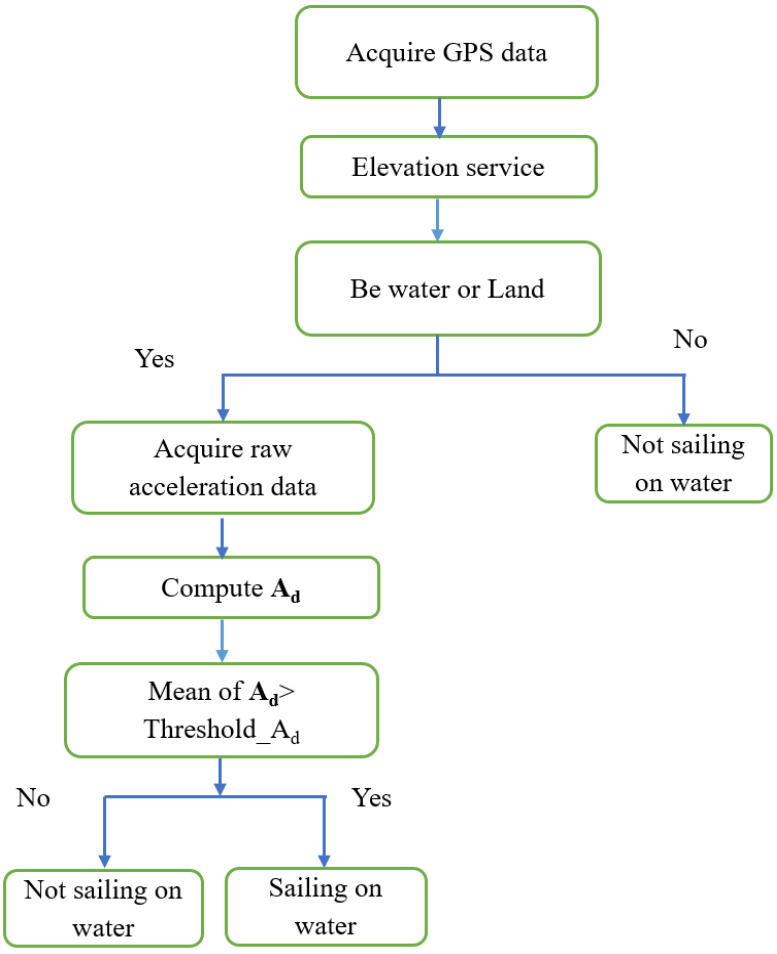


***Figure 1**. Flowchart for the detection of the status "Sailing on Water".*

2.2. Collision warning

In this work, the acceleration data in the phone will be collected and processed to identify collisions that occur when compared to a preset threshold (denoted by Threshold_B_d). Acceleration is known to represent a change of velocity. When there is a collision, the velocity will decrease rapidly to zero. When the acceleration information notifies the state of the collision of the ship, the software will check previous states to see whether it is on sailing to avoid false alarms. False alarms can be the cases when the phone is on the table suddenly dropped to the ground or the phone is in the user's pocket while jogging, falling, etc... When detecting signs of an accident, a warning dialogue box will appear for users to confirm the actual situation. After 60 seconds, if there is no confirmation from the user, the software will treat as an accident and will automatically call relatives to be able to support promptly.

3. RESULTS

Currently, our software is available on the GooglePlay application market as shown in Figure 2. Users can search for products by the following link:

https://play.google.com/store/apps/details?id=vnu.uet.boatsafe

The software has an intuitive interface for easy using. When the application is downloaded, the device asks a series of questions to ask permission from the user. This permission will allow the software to access the phone book, built-in sensors and GPS.

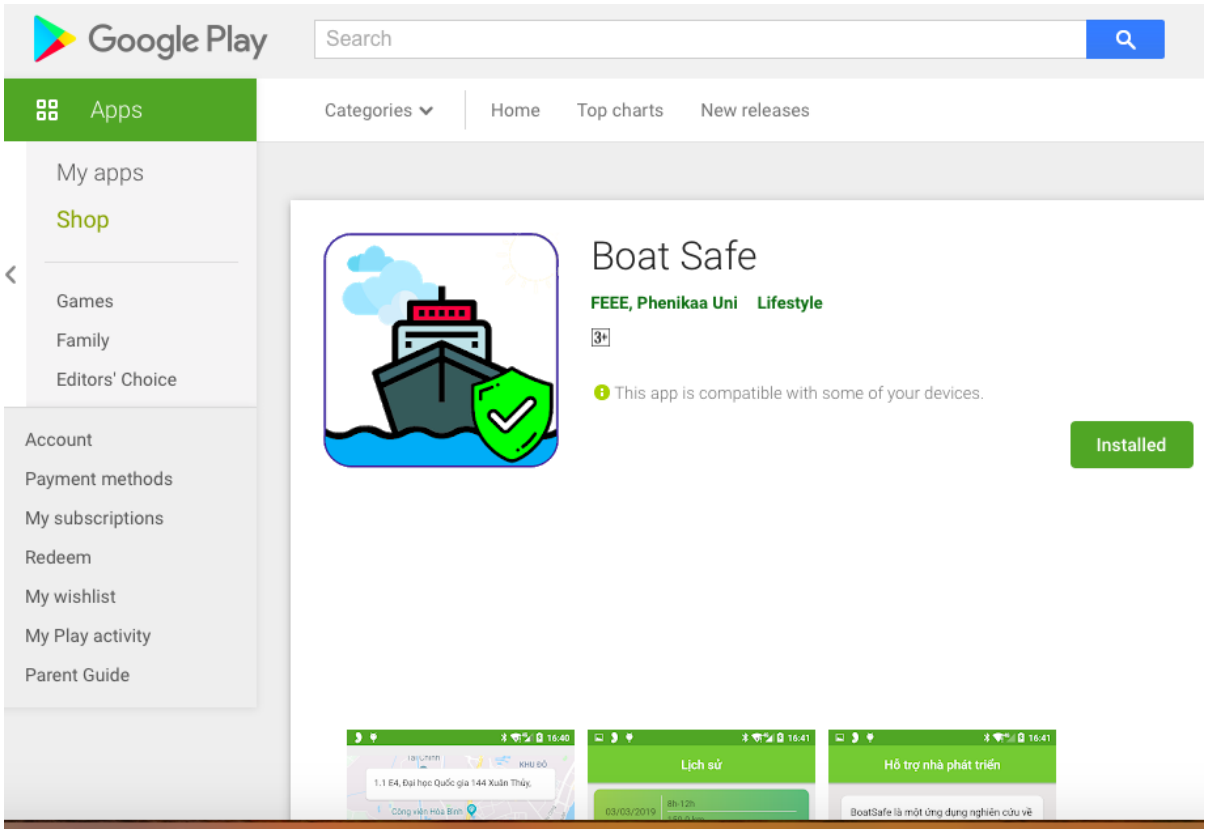


***Figure 2**. Boatsafe software is available on the GooglePlay.*

The information of positions and velocities, of the ship on the screen of the smartphone are shown in Figure 3. If the user saves the history of the sailing, he or she can draw the previous trajectory of the ship (shown in Figure 4). It would be useful for managements.

Figure 3. *The information of positions and velocities on the screen.*

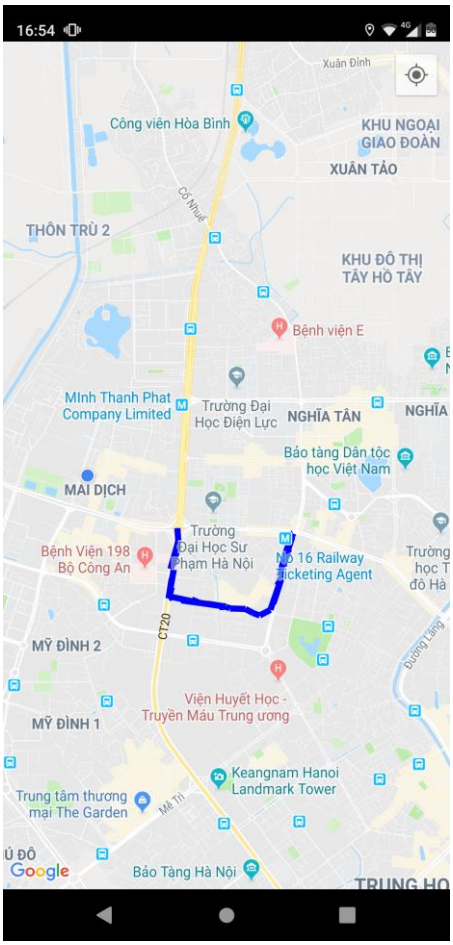


Figure 4. *Users can display the saved trajectory.*

The acceleration data in X, Y, Z directions are shown in Figure 5. There is an event that the absolute acceleration excessed the collision threshold. Consequently, a warning message will appear in the confirmation from the user. If "I'm OK" button is pressed, the warning message will disappear. After 60 seconds, if there is no confirmation from the user, the software will treat as an accident and will automatically call relatives to be able to support promptly.

At present, it can only detect collision when the collision already occurs (Speed to Zero). This feature is needed to improve because in this way the people on a boat may also realize the dangerous situation. In that case, people will actively call relatives or authorities to rescue through the on-board Ship Security Alert System (SSAS). This system is also automatic and may work better than a system on a mobile phone.

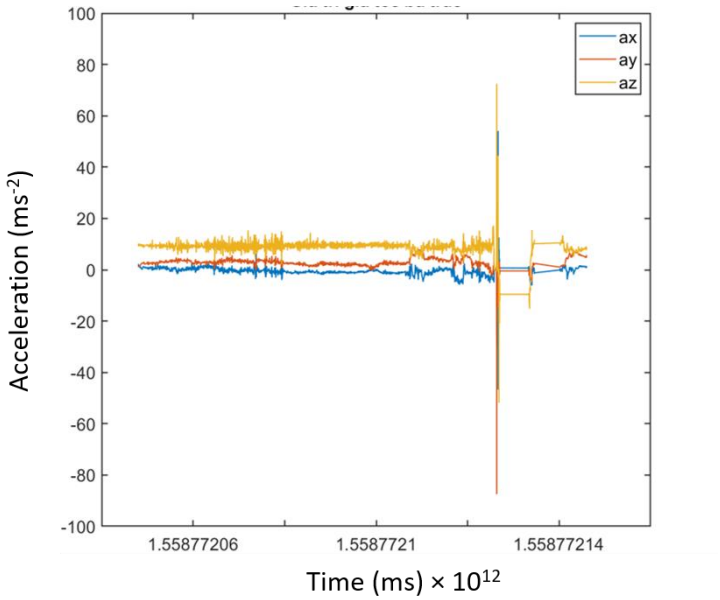


Figure 5. *Acceleration samples which excess the collision threshold.*

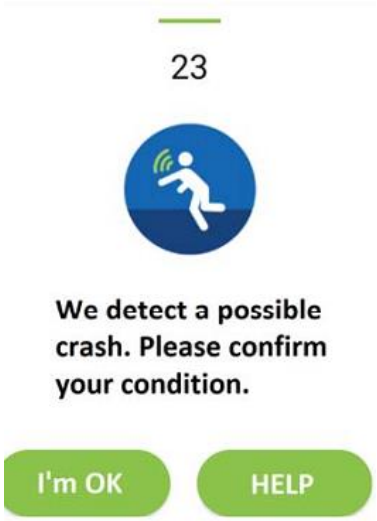


Figure 6. *Warning message.*

4. CONCLUSION

The accuracy of positioning is not better than a built-in modern system with GPS exclusively for boats. However, our software provide more the flexibility and features compared to previous products. Moreover, the software should be tested more with the different inland watercrafts. Even that BoatSafe application is made public, it is still in the development phase. There some features that need to be improved and added in future such as: 1) improve the accuracy of estimation algorithms of positions and velocities; 2) develop on other operating systems such as iOS, Window Phone, etc; 3) empirically verify collision warning; 4) predict the trajectory when the GPS connection lost in a certain of time.

REFERENCES

[1] Luis C. Blancas and M. Baher El-Hifnawi, Facilitating trade through competitive, low-carbon transport: the case for Vietnam's inland and coastal, Worldbank, 2014.

[2] ATG, GPS solution for waterway transport, atgvn.com, access Feb. 10th, 2020. (In Vietnamese)

[3] VietGlobal, Ship locator, vlg.com.vn, access: access Feb. 10th, 2020 (In Vietnamese)

[4] Huy Lộc, Making electronic maps for river ships, Traffic newspapers, 2016, http://www.baogiaothong.vn (In Vietnamese)

[5] Do, D. D., Nguyen, H. V., Tran, N. X., Ta, T. D., Tran, T. D., & Vu, Y. V., Wireless ad hoc network based on global positioning system for marine monitoring, searching and rescuing (MSnR), in Asia-Pacific Microwave Conference, 2011, IEEE, pp. 1510-1513.

[6] Duc-Tan, T., Fortier, P., & Huynh, H. T., Design, simulation, and performance analysis of an INS/GPS system using parallel Kalman filters structure, REV Journal on Electronics and

Communications, 1.2(2011). DOI: http://dx.doi.org/10.21553/rev-jec.15

[7] Van, T. N., Duc, T. C., & Duc-Tan, T., Application of street tracking algorithm in an INS/GPS integrated navigation system, IETE Journal of Research, 61(2015), 251-258. https://doi.org/10.1080/03772063.2015.1019940

[8] Van Thanh, P., Tran, D. T., Nguyen, D. C., Anh, N. D., Dinh, D. N., El-Rabaie, S., & Sandrasegaran, K., Development of a real-time, simple and high-accuracy fall detection system for elderly using 3-DOF accelerometers, Arabian Journal for Science and Engineering, 44(2019), 3329-3342. https://doi.org/10.1007/s13369-018-3496-4

[9] Pham, V. T., Nguyen, T. A., Tran, D. N., Nguyen, D. A., & Tran, D. T., Development of a real time supported program for motorbike drivers using smartphone built-in sensors, International Journal of Engineering and Technology (IJET), 9(2017). 257-270 DOI: 10.21817/ijet/2017/v9i1/170902301.

POSE INVARIANT FACE RECOGNITION SYSTEM

Nguyen Dong Dong[1, 2], Bui Ngoc Dung[1*], Hoang Xuan Tung[3], Tran Nguyen Cac[4]

[1]*University of Transport and Communications, No 3 Cau Giay Street, Hanoi, Vietnam.*

[2]*Samsung Vietnam Mobile R&D Center, No 1 Pham Van Bach Street, Hanoi, Vietnam.*

[3]*VNU University of Engineering and Technology, 144, Xuan Thuy Street, Hanoi, Vietnam.*

[4]*Research and Development Division, Genexis B.V., Netherlands.*

[*]*Corresponding author: dnbui@utc.edu.vn*

Abstract: Face recognition has become popular in security surveillance systems. These systems work well only when the face is in frontal and the system itself has considerable training data. This paper proposes a method of face recognition with various view using Generative Adversarial Networks (GAN). In this method, a pre-trained model of face is used to generate the training data of different pose variations from several frontal faces using GAN. These faces will be used as input of the FaceNet and Support Vector Machine to be trained to perform the task of face recognition. The experiment results demonstrate the accuracy of the proposed method in the case of lacking the training data. This paper focus on the applicability of Generative Adversarial Network in Face Recognition System by applying the GAN-based method in Face Recognition System to increase the number of training samples and detect the faces of different poses that was not possible with the conventional Face Recognition System. The comparison between Face Recognition System using GAN and conventional Face Recognition System was presented.

Keywords: face recognition, generative adversarial network, facenet.

1. INTRODUCTION

Face recognition is a highlight in image processing field, a computer application which is able to automatically recognize human face through a digital image or a video frame. Face recognition system is used in dealing with security problems, detecting and identifying a person in restricted area. A face recognition system commonly includes three main parts: face detection, facial feature extraction and identification of the face by comparing its facial features with stored information in database. Due to the advances in deep learning, the accuracy of face recognition methods is improved significantly. However, face recognition in practice always comes with challenges like pose, illumination, face expression and maybe their combinations. Besides, it requires a great deal of training data for desired accuracy.

There have been many researches trying to solve the problems of the complex face recognition. In [1], Principal Component Analysis (PCA) was used to extract the features from the face and recognize using neural network. This method overcomes the disadvantage of illumination sensitive of PCA [2]. A multi-view deep network has been used to find the

nonlinear discriminant and view invariant representation shared between views [3]. To recognize the pose-invariant face, [4] using GAN to generate the high quality images with 9 different views for training data. Multi-scale local binary pattern features from 27 landmarks were extracted to form the feature vector for pose in variance face recognition [5]. In [6], pose robust feature obtained from the combination of component level and landmark level to overcome the nonlinear intra personal variation [7]. Pose invariance face recognition also be used by Fully-trained Generative Adversarial Networks (FTGAN) for pixel transformation to achieve coarse face transformation, and these faces are refined by key point alignment [7]. Disentangled Representation Learning GAN (DR-GAN) [8] is a method for PIFR (Pose-Invariant Face Recognition) whose purpose is to solve the problem of signficant pose discrepancy between two faces in face recognition, which is not same as the purpose of our proposed system.

In this paper, we propose a new method of face recognition with pose invariance of faces but lack of training images. The system is capable of recognizing faces of different poses and influenced by external agents, while performing facial representations from a fixed posture. Especially, with only a small amount of training data, the system still produces positive recognition results thanks to the ability to re-represent faces from the original data set. To do this, beside the frontal faces, we use GAN to generate fake faces from different poses. This step helps the system recognize the non-frontal faces and increases the training data. These faces were given to the input of FaceNet to extract high quality features, also known as face embedding. Finally, Support Vector Machine (SVM) will be used to train a face identification system. The general framework is shown in figure 1.

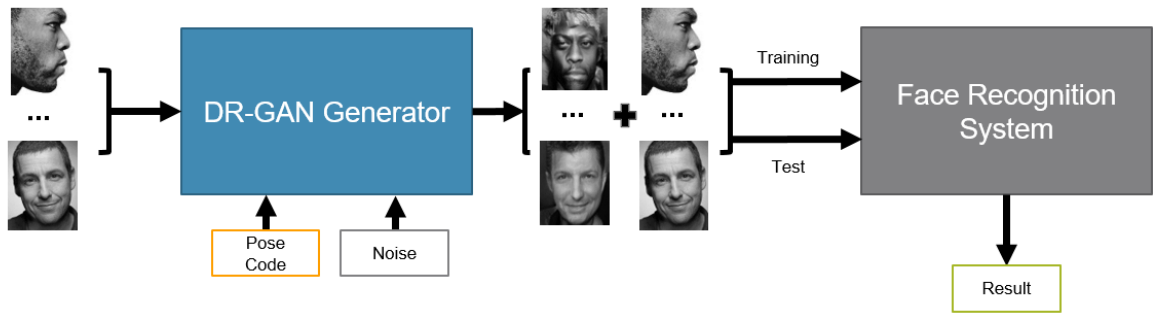


***Figure 1**: DR-GAN Generator works as a face rotator.*

2. METHODOLOGY

Generative Adversarial Networks (GAN) [9] consists of two main components. Generator G is a generative model and discriminator D is a distinguishing model. G takes a probability distribution as input and tries to synthesize a realistic image. D takes two input images, one being the real image from training data and another being fake image generated by G. It tries to determine whether the input is a fake or a real one. D and G are trained

together, D tries to become more efficient in distinguishing between real and fake images and G tries to generate more realistic images that D would not be able to recognize [9].

The process of face recognition using GAN includes 4 steps:

- Generate frontal face images from original training data.
- Face detection.
- Facial feature extraction.
- Face recognition.

Using the frontal and lateral face images, Disentangled Representation Learning GAN (DR-GAN) [8] generator generates extra frontal face images. Training data includes original images and generated images of which a proportion is used for validation. Training images are processed to detect the faces using Haar-Like feature method with AdaBoost and Cascade Classifier algorithms. These detected faces were given to the input of the FaceNet in order to extract high quality facial features. Finally, SVM will be used to train the face identification system.

2.1. Frontal face images generation using DR-GAN.

Input images will be resized to 96x96 and passed into the encoder of generator G_{enc}. Output of G_{enc} is a 320-dimension feature vector $f(x)$. $f(x)$ will be combined with pose code c and noise vector z in order to generate a $(320 + N^p + N^z)$-dimension combined vector which N^p is the total number of pose and N^z is noise vector dimension. A sequence of fractionally-strided convolutions (FConv) in decoder of generator G_{dec} transforms combined vector into a synthetic 96x96-dimension image $\hat{x} = G(x, c, z)$ [8]. The face generation process is shown in figure 2.

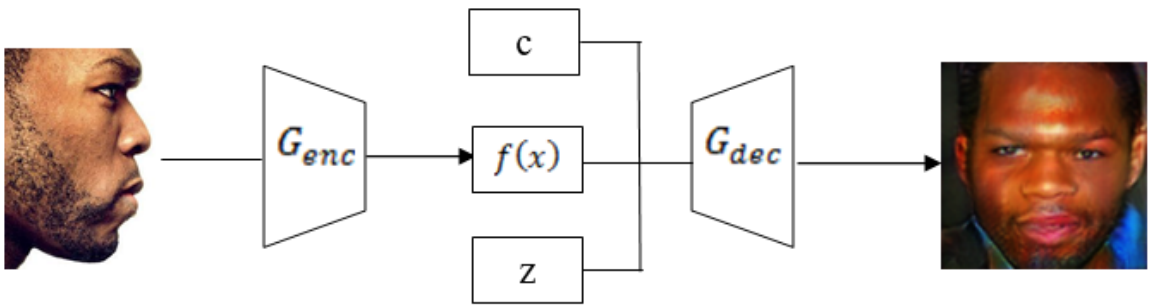


***Figure 2**. Frontal face generation.*

2.2. Facial feature extraction.

In FaceNet, input images are transformed into a set of 128-dimension vectors in which Euclidean distance of similar faces is shortest and vice versa. These vectors are called facial feature vector or embedding vector, which are used for training in the face recognition system.

Face recognition model:

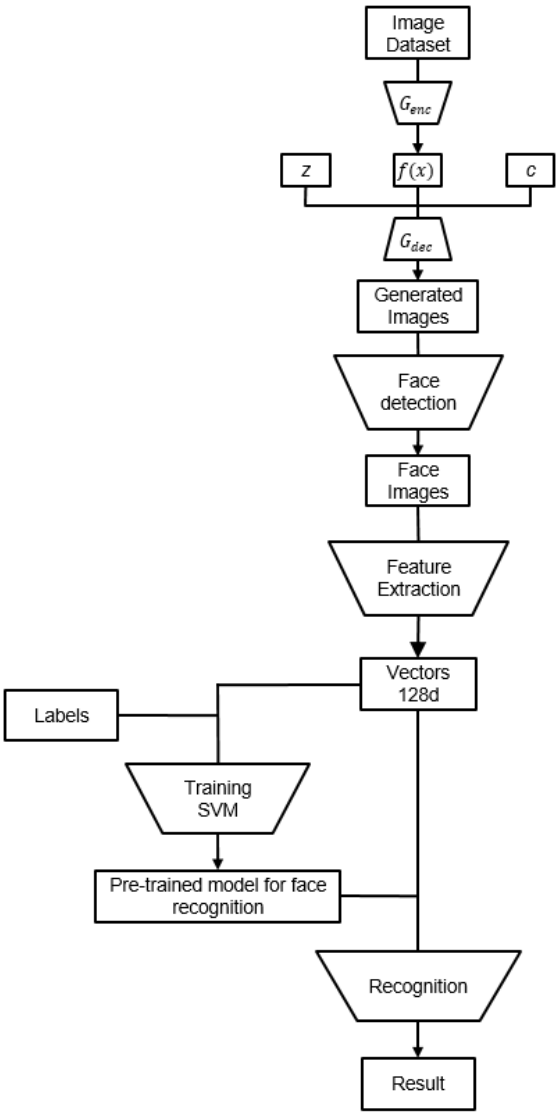


***Figure 3**. Training and recognizing processes*.

Where:

- $Image\ Dataset$: A small data set includes frontal face images and lateral face images of each individual.
- G_{enc} : encoder of *generator* in DR-GAN.
- G_{dec} : decoder of *generator* in DR-GAN.
- $f(x)$: 320-dimension feature vector given by encoder.
- c : pose code.
- z : noise vector.
- $Vectors\ 128d$: 128-dimension facial vector given by FaceNet.

3. EXPERIMENTAL AND RESULTS.

3.1. Data

Figure 4*. Some photos in training data set.*

This research study uses a small sample of data in CFP (Celebrities in Frontal-Profile in the Wild) dataset. The sample consists of 10 frontal-face images and 4 lateral-face images of several celebrities, which will be used as input data. The accuracy of system depends largely on the amount of input data. However, through practical training in a small scale, the accuracy of the proposed face recognition system is relatively positive.

3.2. Image generation using GAN.

With the sample input, Disentangled Representation Learning GAN (DR-GAN) then generates a new set of images that share several features with the input, and recognizable by the face-detecting methods.

DR-GAN generates fake images look like original images of training data.

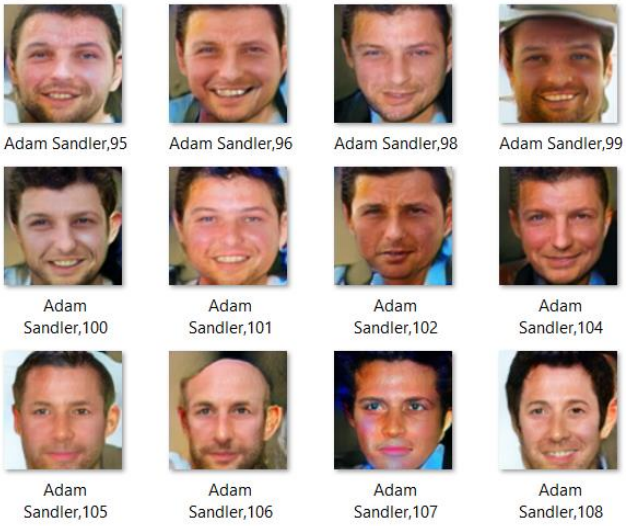


Figure 5. *Generated images.*

In figure 5, the images were generated from Adam Sandler's sample. In general, these generated images have Adam Sandler's facial features. Each image was indexed to name with numerical order for testing purpose.

3.3. Face identification.

The output vectors of Facial feature extraction are then identified by the SVM using pre-trained models.

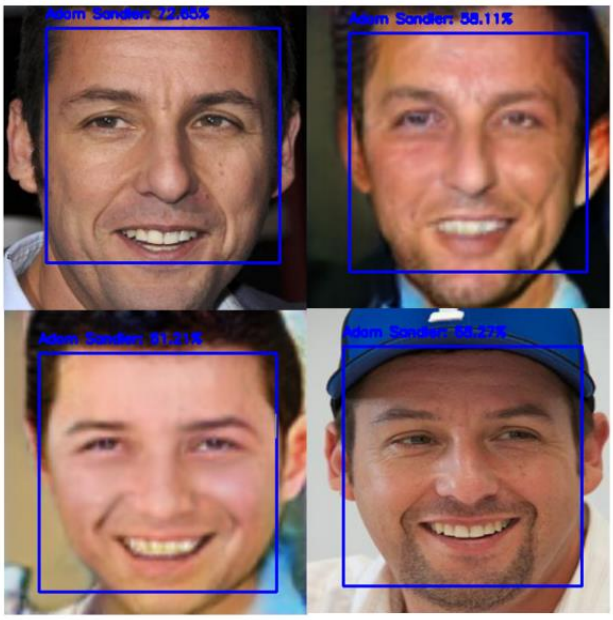


Figure 6. *Results of face identification.*

In figure 5, Adam Sandler's images (both generated and original images) are correctly identified. The bounding box shown face position and the result of identification including name and reliability.

* Result evaluation.

The result of face identification highlights that the proposed system is able to provide a favorable result compared to the conventional system.

With a test sample of 18 images (not included in the training data, including frontal and lateral face images), conventional system is only able to recognize correctly 4 out 18 images (only the frontal ones), with the highest reliability of 90.12%. Whereas, the proposed system using GAN gives better test result with 11 out 18 images (both frontal and lateral) correctly recognized, and the highest reliability is 93.98%.

Figure 7. *The results of the image generation from the training data set and comparison of the identification results between the conventional face recognition system and the proposed one using GAN.*

In figure 7, the training data includes both generated and original images. As the result, conventional face recognition system only is able to identify some individuals by their frontal face. Whereas, the proposed system using GAN can identify someone by both their frontal and lateral faces. Images that are correctly recognized are those with the reliability indicator beneath.

4. CONCLUSION

This paper proposed the face recognition framework that can recognize the face at difference poses. Generative Adversarial Network can act as processing tool that generate the synthesis faces in the case lack of training images and poses. This part helps to solve the problem of the pose difference between two faces by fix the pose and learn the representation of features from the face. The results show our proposed method outperformed the conventional method.

REFERENCES

[1] Mayank Agarwal, Nikunj Jain, Mr. Manish Kumar and Himanshu Agrawal, "Face Recognition Using Eigen Faces and Artificial Neural Network", International Journal of Computer Theory and Engineering, Vol. 2 (4), pp. 1793-8201, August, 2010.

[2] Turk, M., and Pentland, A., "Eigenfaces for recognition", Journal of Cognitive Neuroscience, Vol. 3, pp. 71-86, 1991.

[3] Kan, M., Shan, S., Chen, X.: Multi-view deep network for cross-view classification. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp.4847–4855, 2016.

[4] Avishek Bhattacharjee, Samik Banerjee, and Sukhendu Das, "PosIX-GAN: Generating multiple poses using GAN for Pose-Invariant Face Recognition", ECCV 2018.

[5] Dong, C.; Cao, X.; Fang, W.; Jian, S. Blessing of Dimensionality: High-Dimensional Feature and Its Efficient Compression for Face Verification. In Proceedings of the IEEE Conference on Computer Vision & Pattern Recognition, Portland, OR, USA, 23–28 June 2013, pp. 3025–3032.

[6] Ding, C.; Choi, J.; Tao, D.; Davis, L.S. Multi-Directional Multi-Level Dual-Cross Patterns for Robust Face Recognition. IEEE Trans. Pattern Anal. Mach. Intell. 2016, 38, 518–531.

[7] Weiwei Zhuang, Liang Chen, Chaoqun Hong, Yuxin Liang, and Keshou Wu, "FT-GAN: Face Transformation with Key Points Alignment for Pose-Invariant Face Recognition", electronics, 2019.

[8] Luan Tran, Xi Yin, Xiaoming Liu (2017), Disentangled Representation Learning GAN for Pose-Invariant Face Recognition, Department of Computer Science and Engineering Michigan State University, East Lansing MI 48824.

[9] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative adversarial nets. In NIPS, 2014.

A SIMULATION OF INTELLIGENT TRAFFIC CONTROL USING FUZZY LOGIC UNDER MIXED TRAFFIC CONDITTION

Vuong Xuan Can[1,2,*], Nguyen Hoang Van[2], Vu Trong Thuat[2], Nguyen Thi An[3]

[1]*Southwest Jiaotong University, Chengdu, China*

[2]*University of Transport and Communications, No 3 Cau Giay Street, Hanoi, Vietnam*

[3]*National Economics University, No 207 Giai Phong Street, Hanoi, Vietnam*

[*]*Corresponding author: vuongcan@utc.edu.vn*

Abstract: Recently the large cities of Vietnam, especially Hanoi and Ho Chi Minh City are facing traffic problems, involving traffic congestion, traffic accident, environmental pollution, etc. One of the causes is almost intersections in cities of Vietnam using the pre-timed signal timing plan, which is not suitable for signalized intersections with the high fluctuation of traffic demand and mixed traffic flow dominated by motorcycles. This paper provides a VISSIM-PYTHON-based simulation of intelligent traffic control using the fuzzy logic approach at a signalized intersection under mixed traffic conditions. The simulation results of a typical signalized intersection in Hanoi show that the average delay of the fuzzy logic approach is considerably reduced compared to the existing traffic control plan.

Keywords: simulation, signalized intersections, intelligent traffic control, fuzzy logic, mixed traffic conditions, VISSIM-PYTHON

1. INTRODUCTION

The improvement of traffic signals at intersections is an important key in improving the capacity and safety of an urban road network. The traffic performance of signalized intersections has important impacts on the traffic performance of the whole urban road network. In many cities of Vietnam like Hanoi, Ho Chi Minh City, traffic congestion, traffic safety, air pollution and so on are becoming serious problems, especially at signalized intersections where the delay time and the queue length are rising sharply. One of the main causes leading to those problems is the signal control plan of most signalized intersections based on fixed-time control without optimal cycle length. Moreover, the design specifications for signalized intersections are not detailed enough, so the control plan is usually designed based on the experience of experts and the guidelines of developed countries like Germany, the USA, and France where traffic conditions are quite different from developing countries like Vietnam [1]. Thus, a direct application of the guidelines maybe not appropriate, and impacts on the efficient operation of traffic flow.

The fixed-time signal control methods or called the conventional methods (Webster, 1958[2]; Akcelik, 1981 [3]) may work quite effectively during normal traffic demand, but it is not suitable for signalized intersections with high fluctuation of traffic demand. To overcome the shortcomings of fixed-time control, the intelligent control approaches using intelligent algorithms (e.g., fuzzy logic, genetic algorithm, particle swarm optimization, and so on) are developed in recent years. They have been found to be more cost-effective and

efficient at managing the complex traffic situation. Fuzzy logic (FL), which first was introduced by Zadeh [4-5], has been used widely to develop intelligent control plans at intersections because it allows describing and qualitative modeling of complex systems which have inherent uncertainties that it is not easy to solve using conventional mathematical models [6-7]. Traffic control using the FL can be allowing traffic flow more smoothly, shorter waiting time, and lower stopping percentages at an isolated intersection. The first application of the FL in traffic control was introduced by Pappis and Mamdani [6]. However, most of existing intelligent control, including the FL has been developed for homogeneous traffic conditions. This is quite different from heterogeneous traffic conditions like Vietnam where different vehicle types with a huge motorcycle shared in and compete together for the same road space at intersections. Hence, it's necessary to have an improved control plan which is more suitable for the mixed traffic flow with motorcycle-dominated in developing countries like Vietnam. Besides, traffic signal control using the FL is still new in Vietnam, excepting only a few works that have an interest in simple traffic conditions [8-9]. They don't reflect mixed traffic flow with the domination of motorcycles in Vietnam. Hence, the main objective of this paper is to implement a VISSIM-PYTHON based simulation of signal control plan at an isolated intersection based on fuzzy logic, with reference to the mixed traffic in developing countries like Vietnam.

2. METHODOLOGY

Consider a typical signalized intersection with mixed traffic flow coming from east, south, west, north four approaches. There are left and right turns considered for each approach. To implement intelligent control, assume that some vehicle detectors by traffic cameras in each approach provide meaningful information of actual traffic flow, which includes traffic volume, queue length, occupancy rate and so on. Traffic information in each approach will be determined by vision-based cameras using image processing techniques. The fuzzy-based control determines the duration and sequence that the traffic light should stay in a certain state, before switching to the next state. The principle as follows: 1) Meaningful information of actual traffic flow is collected from vehicle detectors; 2)The traffic information as the fuzzy controller's input, based on the input linguistically determines the duration of green signal as a qualitative input to a traffic controller which controls the signal timing of the traffic light. The principle diagram of the intelligent traffic control at the intersection using fuzzy logic is shown in Figure 1. In this paper, the models are implemented by using VISSIM-PYTHON simulator.

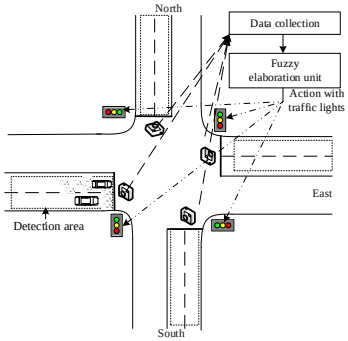


Figure 1. *The principle diagram of the intelligent traffic control fuzzy logic.*

The paper uses two input fuzzy variables about traffic information, including maximum queue length (QL) and arrival to an intersection (V), which collected during the previous cycle, to estimate the green time required for a signal group (stage) during the next cycle. The structure of the proposed fuzzy-based control is shown in Figure 2.

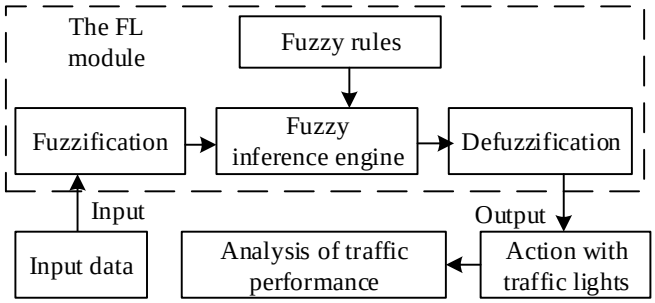


Figure 2. *The structure of the proposed fuzzy-based control.*

The fuzzification is a process that coverts each numerical input into sets of degree of membership by membership functions. Input data could be acquired by camera detector installed at different locations. The detection area of vehicle is placed 120m from the stop line of each approach to record queue length and the number of vehicles. Here, queue length is defined as the distance in metres from the stop line over which vehicles are queuing. Then maximum queue length at the onset of green during is the longest tail of queue irrespective of the lane of each phase in which it occurs, every second. The arrival to intersection is the total amount of vehicles on movements in an approach of a phase. The fuzzy input linguistic variables are very short (VS), short (S), medium (M), long (L), and very long (VL) for maximum queue length (QL); similarly, very small (VS), small (S), medium (M), large (L), and extremely large (VL) for arrivals. Fuzzy membership functions and rules are the essential components in the fuzzy inference engine. Because traffic flow does not change linearly in real time, the Gaussian type membership functions of input and output variables are used in place of triangular membership function. The maximum queue length from 0 to 120m and described by Gaussian fuzzy set with a standard deviation of 3 and the constant of Gaussian membership functions of very short (VS), short (S), medium (M), long (L), and very long (VL) are of 20m, 30m, 40m, 50m and 60m, respectively (see Figure 3). Similarly, for the arrival, its range from 0 to 50 and described by Gaussian fuzzy set with a standard deviation of 1.4 and the constant for Gaussian membership functions of very short (VS), short (S), medium (M), long (L), and very long (VL) are of 5s, 10s, 15s, 20s and 25s, respectively (see Figure 4).

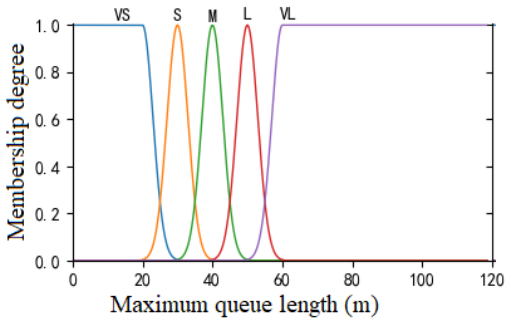


Figure 3. *Membership functions of maximum queue length.*

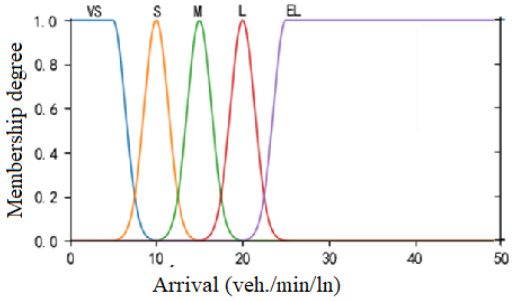


Figure 4. *Membership functions of arrival to intersection.*

The weight of green time (W) is the output parameter of the fuzzy logic and it is divided into 5 ranges, including very low (VL), low (L), medium (M), high (H), very high (VH). All these membership functions are Gaussian type with a standard deviation of 7 and the constant for Gaussian membership functions of very low (VL), low (L), medium (M), high (H), very high (VH) are of 0s, 25s, 50s, 75s and 100s, respectively (see Figure 5).

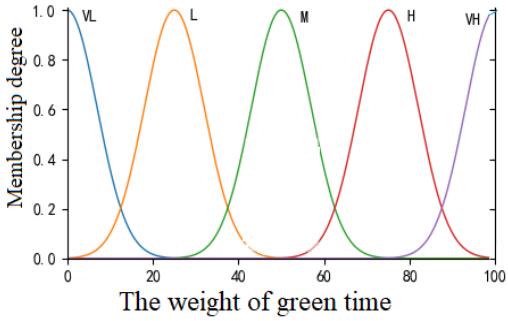


Figure 5. *Membership functions of the weight of green time*

The weight of green time is used to calculate the green time that stages require in the next cycle. According to linear interpolation, we can get the relationship between the green time and the ratio of the amount of green time as follows:

$$g_i = g_{\min,i} + \frac{W_i}{100}.(g_{max,i} - g_{\min,i}) \quad (1)$$

where, g_i is duration of green time of the i-th phase in the next cycle; $g_{max,i}$, $g_{min,i}$ are minimum and maximum green time of the i-th phase, respectively; *Wi* is the weight of green time of the i-th phase. According to the actual traffic experience, green time of a phase cannot be too short, for avoiding vehicles and pedestrians could not pass intersection in time, and it also cannot be too long, considering drivers' mental ability at other phase (i.e., red time for other phases can be long). The minimum green time for each phase is 15.0 seconds, considering little pedestrian volumes and two-lane on each approach [10]. The maximum green time of each phase is determined by calculating the green time at fixed-time control and multiplying by a factor of 1.25 [11].

Fuzzy inference engine is the process of formulating the mapping from a given input to an output using FL. The mapping then provides a basis from which decisions can be made [12]. Fuzzy logic traffic control is designed with rule based on IF-THEN rule statements, as shown in Table 1. For example, if (V = "VS") and (QL = "VS") then (W = "VL").

Table 1. *Fuzzy logic rule base for traffic signal control.*

Arrival (V)	Maximum Queue length (QL)				
	VS	S	M	L	VL
VS	VL	VL	VL	L	L
S	VL	L	L	M	M
M	VL	L	M	M	H
L	L	M	M	H	VH
VL	L	M	H	VH	VH

The defuzzification function in the fuzzy system is a conversion from the final combined fuzzy conclusion into a crisp (non-fuzzy) form that it can implement through various methods (e.g., center of gravity method (COG), mean of maximum method, and height method, etc.). In this study, the OCG is used as a defuzzification technique because it is the most popular defuzzification technique and widely utilized in actual applications.

3. SIMULATION AND RESULTS

A typical intersection, namely Nguyen Chanh-Mac Thai Tong intersection in Hanoi is selected as the research object to tests the efficiency of FL traffic control through series of computational experiments. Nguyen Chanh axis running east-west is a four-lane road without a raised median and Mac Thai Tong axis running north-south is a four-lane road without a raised median, too (see Figure 1). There are 7,593 vehicle movements during a peak hour, including 1,780 cars, 5,594 motorcycles, 22 bicycles, 63 mini-buses, 47 buses, and 87 trucks; The traffic signal control features the pre-timed signal with cycle length of 66s without all-red time. The first phase corresponds to a green time of 31s for the north and south approaches. The second phase corresponds to a green time of 29s for the east and west approaches. The yellow interval is 3 seconds for each phase. After optimizing by the modified Wester method, the green time of phase 1 and phase 2 are 15s and 35s respectively. For details refer to [1].

The FL traffic control interfaces with both the traffic light system and the real-time traffic data through VISSIM micro-simulation implemented on PYTHON (see Figure 6). VISSIM can analyse the operating of urban transportation under various traffic conditions based on time interval and the psycho-physical driver behavior model of driver's behavior developed by Wiedemann (1974)[13]. Recently researches have been calibrated VISSIM to match traffic conditions in heterogeneous traffic conditions [14-16]. To emulate real traffic conditions, the longitudinal and lateral of motorcycles and other significant factors in heterogeneous traffic are calibrated and validated based on the relative speed, desired speed, their distance and so on. In this study, most parameters of motorcycle behavior are calibrated based on [14-15] and [17]. All parameters for driving behavior of four-wheeler vehicles are taken as the default values in VISSIM. All traffic simulations are performed for a 15-minute interval, excluding a warm-up period of 3 minutes to prevent from initial loading effects. Average travel delay (seconds per vehicle) is selected as the performance indicator which is computed by using vehicle trajectory data obtained from traffic simulation.

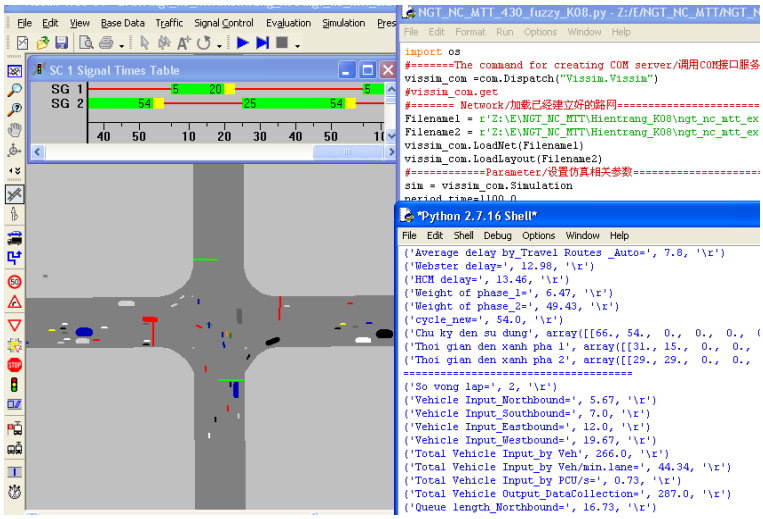


***Figure 6**. Simulation of the FL traffic control by using VISSIM-PYTHON.*

To examine the effectiveness of the FL traffic control, traffic volume has been adjusted from the original flow corresponding to the set of factors K = {0.2, 0.4; 0.6; 0.8; 1,0; 1,2} while keeping the phase number and phase sequence. The set of K coefficients represents the traffic flow scenarios that change with time of day. The average delay at the intersection by using VISSIM-PYTHON is shown in Figure 7.

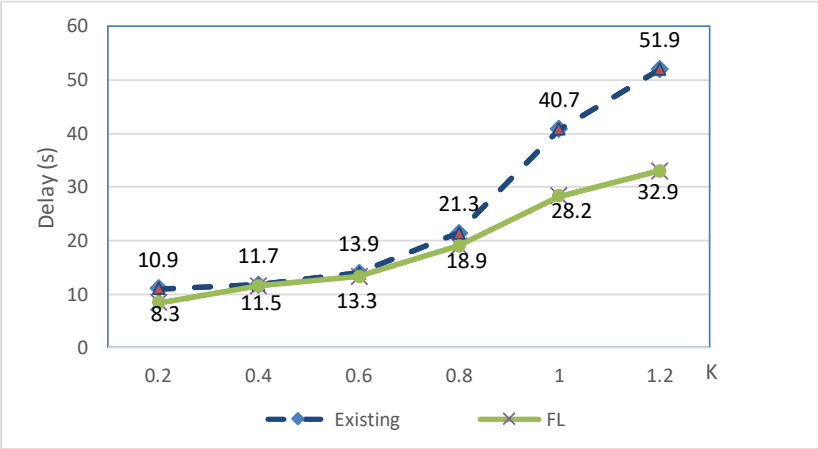


***Figure 7**. Average delay at the intersection.*

Figure 7 shows that FL traffic control considerably reduces the average delay of the vehicles at the intersection, thereby increasing service level of the intersection. There is not a big difference between the FL traffic control and the existing traffic control plan due to low traffic volume for a coefficient of K less than 0.6, but the average delay of the FL traffic control during peak hours is considerably lower than the existing traffic control plan because FL traffic control is more suitable for traffic characteristics.

4. CONCLUSION

The study has implemented the simulation of traffic signal control at intersections based on fuzzy logic under mixed traffic conditions like Vietnam by using VISSIM-

PYTHON. The simulation results show that FL traffic control is more effective than the existing traffic control plan. However, the research has only done simulation for the isolated intersection and has not been tested in the field. The shortcomings should be overcome by the next researches.

ACKNOWLEDGMENT

This research is funded by the Ministry of Education and Training of Vietnam under grant No. B2019-GHA-11.

REFERENCES

[1]. X.-C. Vuong, R.-F. Mou, H.-S. Nguyen, and T.-T. Vu, Signal Timing Optimization of Isolated Intersection for Mixed Traffic Flow in Hanoi City of Vietnam Using VISSIM, in International Conference on Smart Vehicular Technology, Transportation, Communication and Applications, 2018, pp. 133–139.
[2]. F. V Webster, Traffic signal settings, Road Research Technical, London, UK, No. 39, 1958.
[3]. R. Akcelik, Traffic signals: capacity and timing analysis. 1981.
[4]. L. A. Zadeh, Fuzzy logic, neural networks, and soft computing, in Fuzzy Sets, Fuzzy Logic, And Fuzzy Systems: Selected Papers by Lotfi A Zadeh, World Scientific, 1996, pp. 775–782.
[5]. L. A. Zadeh, Fuzzy sets, Inf. Control, 8(3) (1965) 338–353.
[6]. C. P. Pappis and E. H. Mamdani, A fuzzy logic controller for a trafc junction, IEEE Trans. Syst. Man. Cybern., 7(10) (1977) 707–717.
[7]. J. Niittymäki and M. Pursula, Signal control using fuzzy logic, Fuzzy sets Syst., 116 (1) (2000) 11–22.
[8]. N. C. Ngon, Research on Smart Traffic Light System Design, Cantho Univ. Joural Sci., 15b (2010) 56–63 (In Vietnamese).
[9]. H. T. T. Ha, K. A. Tuan, and L. Van Lam, Fuzzy Logic Model for Interlligent Traffic Light Control, in The 9th National Conference on Fundamental and Applied IT Research (FAIR), 2016 (In Vietnamese).
[10]. TRB, Highway Capacity Manual, 2000.
[11]. R. P. Roess, E. S. Prassas, and W. R. McShane, Traffic engineering. Pearson/Prentice Hall, 2004.
[12]. C. Van Pham, Signalized fuzzy logic for diamond interchanges incorporating with fuzzy ramp system, Master thesis, Massey University, 2009.
[13]. PTV.AG, VISSIM 4.30 User Manual, Karlsruhe, Germany, 2007.
[14]. T. H. Van, J.-D. Schmoecker, and S. Fujii, Upgrading from motorbikes to cars: Simulation of current and future traffic conditions in Ho Chi Minh City, in Proceedings of the Eastern Asia Society for Transportation Studies Vol. 7, 2009, p. 335.
[15]. T. Q. Duy, Study for calibrate parameters of car following model which is applied to simulate the motorcycle traffic by traffic simulation method, Master thesis, Ho Chi Minh City University of Technology, 2012 (In Vietnamese).
[16]. S. M. P. Siddharth and G. Ramadurai, Calibration of VISSIM for Indian Heterogeneous Traffic Conditions, Procedia - Social and Behavioral Sciences, 104 (2013) 380–389.
[17]. C. C. Minh, Analysis of motorcycle behavior at midblocks and signalized intersections, PhD thesis, Nagaoka University of Technology, 2007.

THE SOLUTION FORMULATION FOR MONITORING REMOTE PARAMETERS OF RAILWAY EQUIPMENT BY IoT MODEL

Nguyen Van Nghia[1*], Do Viet Dung[1], An Thi Hoai Thu Anh[1]

[1]*University of Transport and Communications, No 3 Cau Giay Street, Hanoi, Vietnam.*

[*]*Corresponding author: nguyennghia.ktd@utc.edu.vn*

Abstract: The paper introduces remote monitoring solutions for locomotives, which applied for the Vietnam Railway industry. The system uses IoT communication model with MQTT communication standard on mobile data network. The theoretical study has been proved by the practical model that has been tested on several locomotives operating on the Vietnam Railway network. Although mobile phone networks along the railway lines have not in good connection, the system receives data from locomotives with high success rates. The practical results verified outstanding advantages of MQTT communication standard and IoT model to the problem of data collection from mobile devices.

Keywords: IoT, MQTT, Locomotive, technical diagnostics, remote control centre.

1. INTRODUCTION

With the operation of the train in both directions on the single track parity, incidents on the route not only slow down the current train, but also cause train stoppage and rescue, which have made the train operation chart change above the whole route for hours later. Therefore, it is necessary to build remote control centres to monitor closely current technical status of locomotive state and predict the state of the machine's imminent fault in an effort to issue a stop command for reasonable locomotive maintenance, improving the operational reliability of the devices.

Technical monitoring systems will allow proactively and independently collect and record all equipment parameters. Without these systems, the technical status only captured through regular maintenance and drivers' reports is incomplete and not really reliable. The drivers on the locomotive cannot observe and record all the parameters in detail because the locomotive system is very complex. When the locomotive installed technical monitoring systems via sensors will be provided complete, accurate and reliable information.

The operation of the monitoring system is moving from offline to online. With the offline system, the data collected on the locomotive will be transferred to the monitoring centre after each working shift through reading the locomotive's data recorders. With the online system, the collected data will be transmitted immediately to the remote monitoring centre via a dedicated wireless data network. Thanks to these data, making judgments of faults, as well as diagnosing technical conditions are carried, so Online monitoring has great efficiency on identifying the status of an upcoming incident, detecting damage on the road, thus avoiding train running obstacles for the entire route.

Depending on the model used, locomotive monitoring centers in the world may be operated by the railway industry or by locomotive maintenance / supply companies [3]. But

the goal is to monitor the technical situation and improve operational quality. In Vietnam, there is currently no interest in online monitoring of technical conditions, networks and online databases about the technical condition of locomotives.

Therefore, the authors have made initial proposals towards building a centralized, remote data monitoring system for locomotives of the Vietnam Railway industry.

The orientation model uses IoT device connection model and wireless communication via mobile communication network and internet. The MQTT (Message Queueing Telemetry Transport) connection model aims to overcome the problems of connection between mobile devices and computers / devices running application programs.

In terms of communications, the world uses the land information system - the specialized locomotive of the railway industry [10], but this system has not still been deployed in Vietnam. Therefore, it is not feasible to apply the same model to Vietnam railway industry. A suitable solution is to use a mobile phone network to connect and transmit data. However, it is necessary to have a solution to ensure the transmission in the condition that the quality of mobile network coverage along the railway is not really guaranteed.

New points of research is designing a model of online technical tracking system for locomotives according to Vietnamese conditions:

- Using MQTT protocol and connecting mobile phone for data transmission

- Using IoT model in designing monitoring applications

2. SETTING THE REMOTE CONTROL METHOD FOR DEVICE RAILWAY IOT MODEL APPLICATION

2.1. General model

The proposed model consists of the following objects (Figure 1).

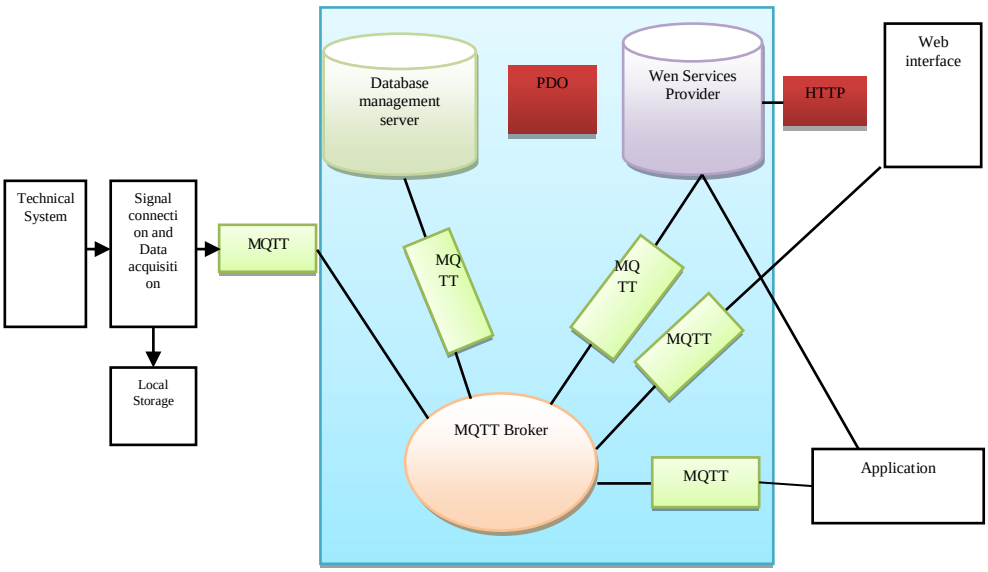


***Figure 1**. Model of remote monitoring system for locomotive technical condition.*

The connection is to collect signals from the locomotive, convert them into digital signals to facilitate the process of storage and transmission. The data are stored on site for technical inspection of locomotives and trains and in case of occurring accidents. Data are transmitted to the control center via data transmission. Data centers receive data and store for technical diagnostics. The diagnostic application uses diagnostic mechanisms to detect special conditions of the locomotive and notify supervisors. Monitoring interface allows supervisors to quickly grasp the technical status of many devices operating in the system and can issue alerts and exchange with drivers.

2.2. The main systems

2.2.1. Connect to the technical system on the locomotive

According to diagnostic requirements, a set of measurement parameters for the device should be built up. The number, measuring range, accuracy, data collection cycle depends on the diagnostic model and diagnostic requirements. This content needs to be completed in combination with designing a diagnose model for locomotive engineering status.

On some types of locomotives, there have been digital control devices (control computers) for the whole locomotive or a functional group of locomotives such as drive control computers, diesel engine control computers... can connect with existing controllers to collect data on the operating status of these devices.

2.2.2. Data transmission

The objective of the data transmission stage is to ensure a safe, reliable and efficient data channel between the equipment and the centre.

Given the current conditions of the Vietnam Railway industry, there is no specialized telecommunications network along the railway, so data transmission based on mobile phone networks is a convenient and effective solution. Using data service of mobile network to connect to data centre through the internet, accepting a small percentage of tracks that are difficult or cannot have instant data, because there is no mobile phone signal.

Use the data service of the mobile network to connect to the data centre through the internet.

Due to the nature of the devices and lines, the connection has some of the following characteristics:

- Communication environment may be interrupted.

- Data transmission is abundant and diverse.

- The transmission path and the source and destination address may be changed during transmission.

Based on the characteristics of the connection, selecting the MQTT communication protocol is the connection model between data devices and data centres. The MQTT protocol enables the management of data sharing between multiple participants and is commonly used for IoT applications. Utilizing the MQTT protocol makes it easy to install a two-way connection between the device and the application via MQTT Broker.

This protocol is standardized by OASIS [2] with the participation of more than 5000 members from more than 600 different organizations and individual members from more than 65 countries and has been approved by ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) in ISO / IEC 20922 [3].

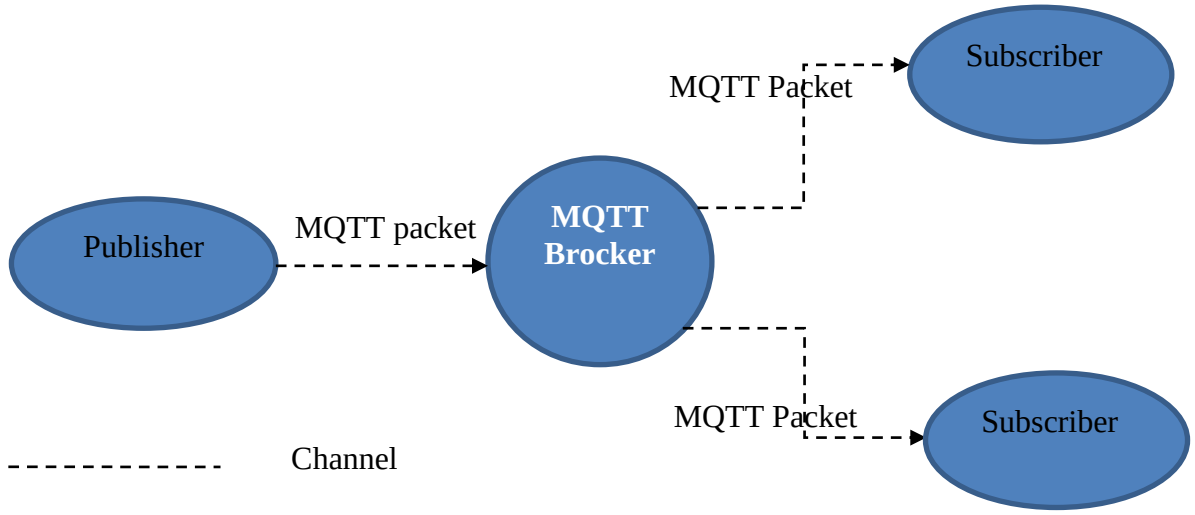


Figure 2. *Sending data mechanism using MQTT protocol.*

The MQTT protocol operates as a publisher / subscriber model as described in Figure 2.

In the operational model of the MQTT protocol, the sender of data is called the Publisher device and the receiver is called the Subscriber device. These devices are also referred to as MQTT clients.

In addition to the Publisher and Subscriber components, the MQTT protocol uses an intermediary device that transfers packets called MQTT Broker. To receive information, the Subscriber devices will send a request to provide information (Subscribe) on a certain topic (topic) to MQTT broker. The subject is a name that distinguishes channels from different Publisher devices and has the form of A / B / C /.../ X / Y tree. When a Publisher device sends notifications providing certain subject information to MQTT broker, the MQTT broker will send that notification to the subscriber devices that requested that topic. The MQTT broker not only plays a role of forwarding notifications from publishers to subscribers based on the topic of the announcement, but also plays other roles such as maintaining connections between publishers and subscribers to MQTT broker, detect abnormal device disconnections, retransmit packets, authenticate and delegate device permissions whenever sending a publish or subscribe notification to the MQTT broker. Currently there are many MQTT Broker developed by brands such as: Mosquitto MQTT, Verner MQTT, Hivemq...

The MQTT protocol is an application-layer protocol, using the transport layer is the protocol that meets three requirements: the packets arrive in order, are not lost and the connection is a two-way connection (like as the characteristics of TCP / IP transport layer protocol). MQTT protocol has 5 outstanding features that make a difference compared to other protocols:

First, the protocol uses the publisher / subscribe model to forward data. This model is suitable for applications that need to forward messages from one sender to multiple recipients.

Secondly, the transmission of packets from Publisher to Subscribers is not related to the content of the packet.

Third, the MQTT protocol divides three levels of service quality for the process of delivering packets from Publisher to Subscriber. With the type of service that transmits at most once, the notification from Publisher will be transmitted to the Subscriber, but if there is a loss of notification during the transmission, there will be no mechanism for retrieving the notification. With the type of service transmitted at least once, the notification from the publisher will be guaranteed to be transmitted to the subscriber, yet the notice may be sent to the subscriber more than once due to the resending of the notification when the confirmation packet is lost. With the correct type of transmission service once, a notification from Publisher will be sent to the Subscriber exactly once.

Fourth, the header size of MQTT packets is very small (fixed fields are 2 bytes in size) compared to other application layer protocols. This makes it possible to operate in network environments with low transfer rates.

Fifth, the notification mechanism when an abnormal connection occurs from a publisher or subscriber device.

MQTT is currently widely utilized in various IoT applications such as Amazon IoT, Microsoft Azure IoT Hub, Facebook Messenger chat application, etc.

2.2.3. Data collection and storage centre

Need to meet the number of connected devices and data retrieval requirements. Data centres receive data from the device and store them into convenient records for lookup. Data centres are often based on platforms like Microsoft Sql Server, Oracle, Mysql.

2.2.4. The diagnostic application

Using diagnostic mechanisms to detect special conditions of the locomotive and notify supervisors.

2.2.5. Monitoring interface

Monitoring interface allows supervisors to quickly grasp the technical status of many devices operating in the system and can issue alerts and exchange with drivers.

The solution carried out in the authors' research is to choose the connection solution and organization of application data centre according to IoT model to promote the strengths of this model.

3. SETTING IOT FOR REMOTE MONITORING OF THE LOCOMOTIVE TECHNICAL STATUS

Based on the above analysis, the authors have designed a model of the technical situation monitoring system on the locomotive. The model aims to demonstrate the system's

ability to perform and its effectiveness in practice.

3.1. Technical connection system

Carrying out a survey on the data structure of technical systems:

1. The communication structures are currently used on locomotives of the Vietnam railway industry

- D19 locomotive: Using the CAN protocol to communicate between the components of a locomotive computer and a diagnostic device [5]

- D20 locomotive: Using MVB (Multifunction Vehicle Bus) protocol allows connection of multi-function devices on vehicles [4], [6]

- Speed measuring device DT04.4: Using RS-485 communication port to connect external devices. [2][5]

In general, these protocols are multipoint protocols, which allow a new device to join the network and receive / send data in the network. However, these are data networks meant for internal control, equipment suppliers do not have official documents on the specifications and meanings of data exchanged. Therefore, more research is needed to be able to extract data from these sources safely and effectively.

2. Sensors receive signals on the locomotive

The locomotive uses various types of sensors to collect data on technical conditions. Vietnam railway uses a variety of locomotives, each of which has a different design and different sensor, measurement and control systems. Generally, information from sensors is processed as follows:

- Sensors receive signals, lead to protection relays and display clocks: Used with old locomotives, without digital control center or independent protective devices, simple parameters do not need coordination between devices

- Sensors receive signals taken to the central control computer to coordinate control, monitoring and alerting the operator: Used with new locomotives, with a digital control center (D19 type machine, D20).

The current sensor system covers all the main information of the technical system on the locomotive. However, the interventions to acquire still need a lot of research time to be able to obtain useful information without affecting the ability of the locomotive to operate safely. Currently, the authors have gained locomotive's operating safety monitoring information by temporarily connecting to the BUS data of DT04.4 monitoring system because this system has been allowed to use widely and consistently on different types of locomotives.

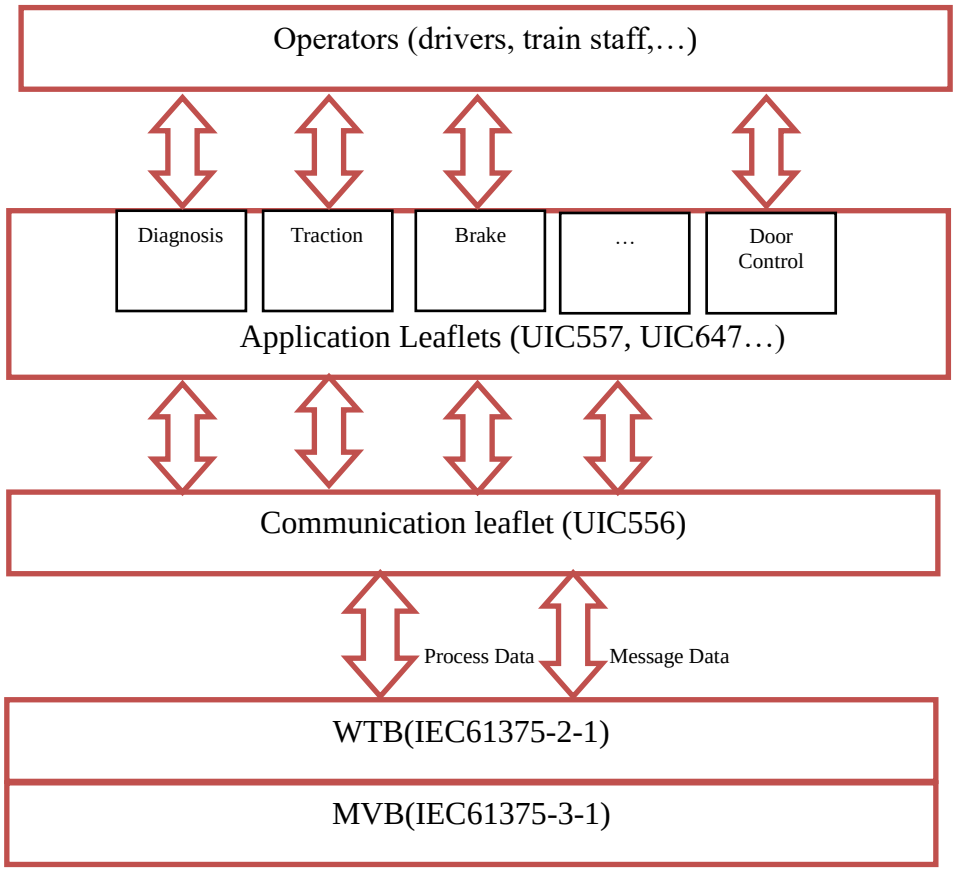


Figure 3. *Data transmission structures used on D20E locomotive.*

Signals that have been measured and used for monitoring and control purposes on the locomotive, are able to pair and receive signals on a basis that does not affect the operation of the device.

It is necessary for unavailable signals to arrange sensors to collect data as required.

For the convenience of the study, the data are taken from the velocity and pressure recorder gauges of DT04.4 locomotive, all of which were installed on some locomotives instead of the locomotive speedometer. The device collects a number of main parameters related to safety and operation of the locomotive such as: locomotive speed, brake pressure, voltage, traction current, locomotive GPS position. Redundancy connects according to CAN standard to the locomotive computer.

3.2. The transmission system

The connection uses 3G mobile network communication module. Connection circuit is designed into a card installed inside the data recording device (Figure 4).

The connection is made via an intermediary communication server using the MQTT protocol. MQTT Broker uses open source Mosquitto software on Linux operating system version Ubuntu 14.

Server used with database with configuration as table 1.

(a)

(b)

Figure 4. *DT0.4 (a) and MQTT connection card (b).*

Communication organization is done through the definition of topics as shown in Figure 5.

- On the data source side: Each device is an MQTT Publisher.

Topics are defined by structure: *data/<device ID>/<context>*

Use protocol one-one to ensure exactly one packet is transmitted to a safe place.

Devices also register the subscriber topic *config/<device name>/<context>* to receive data in the opposite direction.

Table 1. *Parameter of MQTT & Mysql server.*

Parameter	Value
CPU	2 x 6 Core
RAM	32GB
HDD	4 x 300GB
Lan	2 x 1GB
Plattform	1U

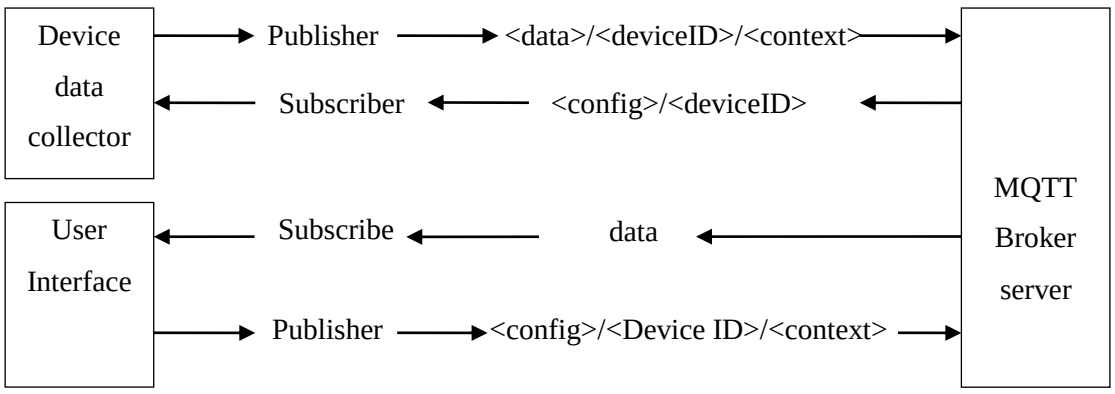


Figure 5. *Data connection model via MQTT server.*

- On the data acquisition application side:

Subscribe the topic *data* to receive data from the device.

Publish the topic: *config/<device ID>/<context>* to send data to the device.

3.3. Database

Using the mysql database version 5.7 running on the Linux operating system version Ubuntu (14).

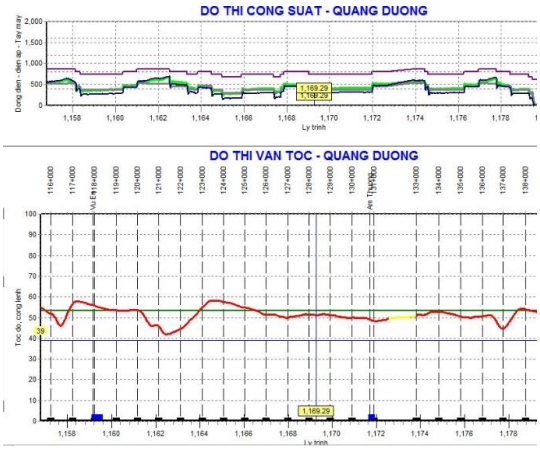
An application is set to subscriber to receive data from the device, checking the correctness and storing it in the Database.

3.4. Query application and data analysis

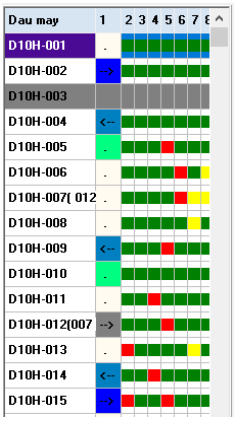
Installing an application for querying and analysing data running on a PC and on the Web. The equipment installed on some locomotives (Figure 6.c) has run stably, and long-term since June 2019. The application named LssClientshas the following features:

- Querying current information of the devices. (Figure 6.a)
- Querying technical data of the devices over time. (Figure 6.b)
- Summarying of technical status of devices in the form of color palettes. (Figure 6.c)
- Monitoring the location and status of the locomotive in the form of a map. (Figure 6.d)

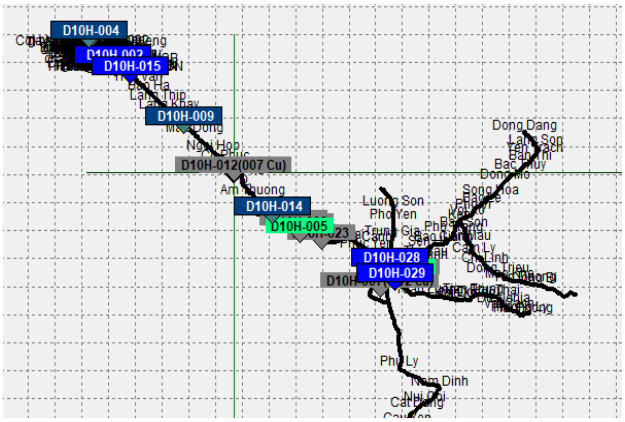
a. Online status



b. Device specs over time



c. Summary of equipment status



d. Location - device status

Figure 6. *Results of monitoring and querying device technical status.*

3.5. Results

The system has been tested on a proposed model.: The device has been installed on locomotives of Vietnam railway industry (Figure 6). From device to control center using wireless connection via mobile data network. Data is transferred to the center on a 2-second cycle. Data is transmitted / received continuously at a low packet error rate. Figure 7 shows the results of checking transmitted data after analyzing the MQTT packet received.

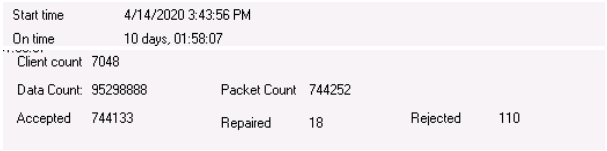


Figure 7. *Results of tracking the successful transmission packet rate.*

Theoretical studies have been verified by experimental models that have been tested on a number of locomotives operating on the Vietnam Railway network. In the condition that Mobile phone networks along the railway lines have the poor connection quality, the system receives data from locomotives with high success rates. The result proves the suitability of MQTT communication standard and IoT model for the problem of data collection from mobile devices.

4. CONCLUSION

In this study, the work of authors has achieved following contributions:

- Proposing a suitable model for the locomotive's engineering status monitoring system based on the structure, protocol and working principles of the IoT model.

- The basic components of the technical monitoring system of the locomotive engineering monitoring system were selected, including a data collection system connected to the locomotive's technical system using communication connections: RS485, CAN; Transmission solution using MQTT protocol.

- Designing a test model using MQTT protocol with the components: Server MQTT Mosquitto; Mysql Database, Subscriber application receiving and analysing data, LssClient query and analysis application.

The designed model operates stably, meeting the requirements, showing the suitability of the IoT solution and MQTT protocol for the problem of collecting diagnostic data. However, in order to set up a diagnostic processing centre, more research is needed on the locomotive engineering diagnostic problem.

REFERENCES

[1]. Andrew Banks and Rahul Gupta, http://docs.oasisopen.org/mqtt/mqtt/v3.1.1/csprd02/mqtt-v3.1.1-csprd02.pdf, OASIS Standard, 2014, pp. 1- 2.

[2]. ISO/IEC JTC1, Information technology - Message Queuing Telemetry Transport (MQTT) v3.1.1, ISO/IEC 20922, 2016, pp. iii.

[3]. Jianghua Feng, Xiangyang Lu, Weifeng Yang, Jun Tang, Survey of Development and Application of Train Communicaition Network, Proceedings of the 2015 International Conference on Eletrical and Information Technologies for Rail Transportation, 2015, pp.843-854.

[4]. Đỗ Việt Dũng, Nguyễn Văn Nghĩa, Nghiên cứu đề xuất mô hình thiết bị đo, hiển thị, lưu trữ dữ liệu vận hành cho đầu máy trên đường sắt Việt Nam, Tạp chí Giao thông Vận tải, ISSN 1859-2724, Số 8/2012, trang 31-33.

[5]. Đỗ Việt Dũng, Nguyễn Văn Nghĩa, Nghiên cứu thiết kế, chế tạo thiết bị đo, ghi tốc độ và một số tham số vận hành cho đầu máy xe lửa D14E, Tạp chí khoa học Giao thông Vận tải, ISSN 1859-2724, số đặc biệt, trang 286-291 (10/2013).

[6]. Vietnam Railways, "Tài liệu kỹ thuật đầu máy D19", 1990.

[7]. Vietnam Railways, "Tài liệu kỹ thuật đầu máy D20", 2006.

[8]. IEC WG43. Electric railway equipment-train communication network-part 2–4: application profile [EB/OL]. www.iec.ch.

[9]. Urabe H (2012) High data rate ground-to-train free-space optical communication system. Opt Eng 51(3):031204:1–9.

[10]. Thaduri, A., Galar, D., & Kumar, U. (2015). Railway assets: A potential domain for big data analytics. Procedia Computer Science, 53, 457-467.

APPLICATIONS OF MULTI AGENT SYSTEMS IN URBAN TRAFFIC REGULATION

Huan Vu[1,2*]

[1]*Université de Lyon, CNRS, Université Lyon 1, LIRIS, UMR 5205, Lyon 69622, France*

[2]*University of Transport and Communications, No 3 Cau Giay Street, Hanoi, Vietnam.*

[*]*Corresponding author: huan.vu@liris.cnrs.fr, huan.vu@utc.edu.vn*

Abstract: In recent decades, agent-based modelling is an approach that has been widely used to solve problems that consist of multiple decentralised individual agents that interact with others. This technique enables the possibility to create and simulate behaviour of different types of agents and the impact of their actions in the environment. Transportation problems such as traffic modelling, intersection management or even parking allocation represent an important field of application. This paper gives a brief overview of several works that have been conducted in the multi-agent community, especially those that deal with regulation or coordination between vehicles in the urban area on different scales.

Keywords: multi-agent systems, agent modelling, traffic regulation, intersection management.

1. INTRODUCTION

With the growth in urbanisation and ownership of vehicles, most major cities around the world suffer from high rates of traffic congestion, with a significant impact on the economy and human well-being. In the US alone, urban congestion costs 8.8 billion hours of travel delay, and 3.3 billion gallons (about 12.5 billion litres) of wasted fuel per year [25]. In addition, pollution due to petrol and diesel cars at stand still at major traffic intersections can rise to more than 20 times than in normal free flow traffic conditions [14]. Several attempts have been made to address this problem for the last decades. Solutions vary from the new designs of lanes, intersections and roundabouts to the optimisation of traffic light plans and traffic flow. With the arrival of new technologies that are developed in recent years, vehicles are now able to use intelligent devices on board (e.g. sensors, communication devices, autopilot systems). These devices enable the possibility to conduct new approaches to coordinate vehicles in urban traffic.

Multi-agent system is a field of distributed artificial intelligence. The system is often composed of multiple intelligent agents, which can take autonomous actions for themselves while interacting with the other agents and their environment. Various issues related to the regulation of urban traffic have been identified by the multi-agent community. Indeed, using agent technology for traffic regulation makes it possible to introduce different types of strategy, with the agents (which are regularly the vehicles) being either fully cooperative (i.e. they cooperate in order to improve the overall performance), partially cooperative or selfish. We overview here several works, in their diversity, the way that they deal with regulation or

coordination between vehicles, on the scale of an intersection or a wider area. In the first case, coordination and regulation at an intersection can be considered as a problem independent of the question of global regulation on a network, and the term "isolated" intersection means that the rest of the network is not taken in consideration. In the latter case, coordination takes place between intersections in order to achieve more coherent network-wide regulation.

The remainder of this paper is organised as follows: Section 2 gives a brief overview of multi-agent systems. Section 3 surveys several approaches based on multi-agent systems for traffic regulation, ranging from the traffic regulation on an isolated intersection to coordination problem for several intersections. We then discuss in Section 4 the differences, the limits of some notable ones, and highlight potential challenges. Section 5 finally concludes the paper.

2. MULTI AGENT SYSTEMS

In distributed artificial intelligence, multi-agent systems represent a sub-field that grows rapidly because of its flexibility in modelling distributed problems. Unlike traditional problem solving methods, agent-based solution consists of modelling the entities in the environment as intelligent agents. Usually, these agents are assumed to have several capabilities such as: to perceive their environment and build local knowledge, to respond to changes to satisfy their objectives, to take initiatives toward the local or global goal and to interact with other agents. Agent-based technology reflects closely to the real-life scenario, especially with the recent development of internet-of-things, where each entity in the system has the capabilities of an intelligent agents and problems are now able to be solved in a distributed manner, enhancing privacy and robustness.

Multi-agent systems have been used in various applications in different domains of artificial intelligence. The applications range from classic problems like telecommunications, traffic control to more modern problems like smart grid or robot coordination problem. In the next section, we will present several multi-agent applications for traffic regulation problems on isolated intersection and on a network of intersections.

3. MULTI AGENT SYSTEMS FOR TRAFFIC REGULATION

3.1 Regulation on an isolated intersection

On an isolated intersection, various coordination problems and various approaches can be considered. Most of them concern real-time traffic regulation, and how the right-of-way is allocated to the vehicles (i.e. giving each vehicle a dedicated time to enter the conflict zone). Some of these approaches imply a regulation agent that performs the regulation alone, others imply inter-vehicular coordination.

A first approach is tackled in [31]. In this paper, vehicles communicate their information to other vehicles in order to coordinate on the intersection. The vehicles have various trajectories that intersect at conflict points. To perform a right-of-way allocation and provide a crossing date to each vehicle, the agents' behaviour is based on a collaboration scheme. Without any coordination, the vehicles have conflicts. With a basic collision avoidance, the vehicles' crossing dates are delayed, one at a time, to avoid these conflicts. In

the collaboration scheme, the vehicles can change the order in which they are delayed in order to minimise the accumulated delay of the vehicles. However, the authors do not provide details about the interaction mechanism performing this minimisation.

[3] uses the notion of fairness for traffic regulation, by proposing a control policy for intersections based on the history of the vehicles. This policy reduces the variance of the total time spent by the vehicles waiting at red lights during their journeys. Each intersection has a controller able to produce several traffic light patterns. A traffic light pattern is a combination of green and red lights duration for each approach, that avoids conflicts. The controller uses various score functions based on efficiency and fairness to evaluate each possible regulation pattern. The efficiency and the fairness of each pattern are evaluated, for various grid sizes and various traffic loads.

Some of the works on the isolated intersection concern vehicle coordination, and others concern the intersection regulation. AIM (Autonomous Intersection Management) aims to coordinate autonomous vehicles at an intersection. Coordinating these vehicles implies granting the right-of-way to the vehicles, so AIM also performs intersection regulation based on the vehicle information. The following articles treat the subject of AIM.

In [11], K. Dresner and P. Stone propose a right-of-way awarding mechanism based on reservations for autonomous vehicles. It relies on a policy called FCFS (First Come First Served), granting the right-of-way to each vehicle requesting, as quickly as possible. This mechanism allows to take into account human drivers [10] by using a classical traffic light policy for human drivers, and giving the right-of-way on red lights to automatic vehicles using the FCFS policy. Although this mechanism can accommodate human drivers, its main benefits are derived from the FCFS policy and the presence of autonomous vehicles.

[15] aims to perform coordination between vehicles approaching an intersection by constructing a priority (oriented) graph. This work proposes a characterisation of feasible priorities using a priority graph. The authors suggest, the use of heuristics to build such graphs in an optimal way at a future stage of this work.

In [30], the different trajectories of the vehicles in the intersection are called streams. For example, all the vehicles coming from the south and going to the west form a stream. Groups of streams are formed, such that groups do not intersect in conflict points. Such streams are called "compatible".

[23] and [24] propose mechanisms to take into account the different valuations of time reduction for the drivers (for example, one minute is more important for a driver being late for a job interview than for a driver driving home from work). In these mechanisms, each vehicle has a budget and can buy or sell time slots. In [23], an auction mechanism called ITSA (Initial Time Slot Auction) is proposed: while joining the neighbourhood of an intersection every vehicle has the ability to bid a part of its budget in order to get the first available time slot and thus cross the intersection. In [24], another mechanism is proposed: TSE (Time Slot Exchange). With TSE, vehicles can trade their respective time slots for credits. A hurried driver will be able to spend what he saved to gain time; other drivers will earn credits. A brokerage agent manages these exchanges according to the demands of each driver. %This time slots are traded with negotiation between agents.

[27] also proposes a market-based approach for AIM. While choosing their itineraries,

drivers are likely to choose the shortest path according to the estimated travel time of each path. In this model, the drivers have to purchase reservations from the intersection managers in order to cross the intersections. This reservation system provides incentives for drivers to explore alternative paths. In this mechanism, each intersection manager has to determine its reservation fare in order to maximise its profit. With few vehicles the intersection manager would earn a low profit, but with numerous vehicles it would actually lose profits because of congestion, so it has to adjust the reservation fare to get an average number of vehicles. In order to perform a relevant joint action in fare adjusting, intersection managers use Q-learning.

[26] proposes a synchronisation-based control to manage traffic at an intersection. In this work, the inner area of the intersection is managed by a control agent. This agent uses an alternating principle in order to determine right-of-way for vehicles from different directions through the intersection. The main idea is to compute the speed profile for each vehicle (i.e. acceleration and deceleration) so that it arrives at the intersection at the assigned time and speed.

In [12], the optimal scheduling for CAVs is dealt with using Mixed Integer Linear Program (MILP). In this mechanism, the constraints that vehicles have to follow in order to build a plan are formalised. These constraints are calculated based on the speed limit and maximum acceleration, as well as the safety gap between vehicles. The gaps are categorised into: gap between vehicles on the same trajectory and gap between vehicles having conflicting movements. Finally, after identifying all the constraints, the mechanism uses MILP to solve the formulated problem, in order to find the optimal solution.

Recently, [28] introduces a novel solution where the regulation problem is modelled as a Distributed Constraints Optimisation Problem (DCOP). DCOP is a technique that has been widely used in the multi-agent community to deal with cooperation problems such as disaster response, task allocation and meeting scheduling. In this work, a DCOP model is proposed and evaluated. The issues related such as computational time, failure of DCOP algorithms have also been addressed.

3.2 Other inter-vehicular coordination problems

Regulation and right-of-way allocation is not the only coordination problem existing at an intersection. In the following approaches, vehicles use vehicular communication to perform a real-time coordination for a pre-existing regulation policy.

In [6], coordination between vehicles is used for traffic simulation. By modelling the intersection problem as a 2-player game where the players are the vehicles, and then as a n-player game, it's possible to simulate realistic human behaviours when a priority rule already exists. The moves of the players are "go" and "stop", and the payoff matrix is built by the players, allowing them to choose the most relevant behaviour. However, the complexity of this method is high and it is difficult to use it for more than a few vehicles.

A common coordination problem for vehicles on an isolated intersection is Collision Avoidance (CA) [7][16]. CA consists in adjusting the speed of autonomous vehicles approaching an intersection in order to avoid collisions, and the solution to this problem often involves the use of mechanical equations to find the appropriate speed for each vehicle. The CA therefore does not concern the policy of regulation although the adjustment of the speed

of the vehicles can affect their admission date in the intersection. However, this aspect is not tackled in detail. For example, [7] states that the approach to the admission order of vehicles can be seen as a simple rule of priority in which the first agent in the sequence has the advantage of keeping its desired motion profile.

[20] presents the right of way as a nonlinear constrained optimisation problem. In this method, the vehicles adjust their speed under various constraints (maximum acceleration and deceleration, maximum and minimum speed, minimum headway distance) in order to minimise the length of the overlapping trajectories of the vehicles in the conflict zone. The computation is performed in a centralised way by an Intersection Control Agent.

3.3 Coordinated regulation for several intersections

The following works are based on coordination on the scale of several intersections. Allowing a larger scale coordination provides a better efficiency of the network, for example with green waves formation. A green wave is a phenomenon consisting in coordinating traffic lights in such a way that a group of vehicles can pass through a succession of green lights, reducing the time loss caused by stop-and-go traffic.

[13] proposes a hierarchical multi-agent model for traffic regulation. In this model, Local Traffic Agents (LTA) perform a regulation at the intersection scale using sensory data. At a larger scale, an Information Traffic Agent stores information about the state of each intersection. At an intermediate scale, Coordinator Traffic Agents (CTA) monitor the intersections of an area to provide information to LTAs about the state of their neighbours, particularly congestion, allowing the LTAs to adjust their behaviours and take into account larger scale information and goals.

In [19], groups of compatible streams are called "signal groups" and are represented by agents. On the intersection scale, each signal group negotiates with the others to get green light or green time extension according to the size of the queues for each signal group. Fuzzy logic is used to determine whether queues have to be considered as short or long in a non-boolean way. On the network scale, intersections are able to exchange their traffic and control data. This allows signal groups to take into account the neighbours' control decision to get green extensions and cause green waves.

[5] represents traffic control as a stochastic game. Traffic has various possible states, and various possible traffic policies are possible actions to be performed by an agent representing the intersection controller. A distributed Q-learning is performed to learn and apply the best traffic policy for each traffic state.

[1] proposes a holonic multi-agent system for traffic signals control. A holonic system is a multi-level system in which each "holon" is made of lower level holons, or atomic agents at the lowest level. In this model, the atomic agent is a signal controller for a single intersection and performs a local regulation based on a Q-learning. Higher level holons represent areas of the network, and their role is to restrict the action space of their sub-holons by giving them abstract actions to perform. Super-holons and sub-holons perform a common Q-learning and each level updates its own policy.

[17] shows how an optimisation between several intersections is possible in an AIM context. On an individual scale, itinerary communication allows each intersection manager to produce an estimation of the crossing time for the vehicles. Then, this crossing time is given

to the vehicles, allowing these to change their itineraries with realistic estimations of their travel times. On a larger scale, this work addresses Braess' Paradox, whereby opening additional travel options for the vehicles reduces the efficiency of all vehicles in the system. Indeed, allowing the vehicles to perform a dynamic itinerary choice with self-interested goals leads to this suboptimal Nash equilibrium in which Braess' paradox occurs. Using dynamic lane reversal, the topology of the network is dynamically changed and avoids Braess' Paradox.

Some methods are based on drivers' behaviour, others on traffic light control. [4] discusses the co-adaptation of vehicles and traffic controllers. Various experiments are made, where only the drivers adapt themselves or only the controllers adapt themselves, or both adapt. They conclude that co-adaptation leads to traffic improvements, especially in large-scale situations involving hundreds of vehicles.

[21] presents an approach for congestion management in CSINs (Complex Self-Interested Networks) using negotiation between agents. The network is represented by a graph and divided into subgraphs, called "worlds". This division of the graph is based on graph properties. The main problem is divided into sub-problems, easier to solve, and the agents negotiate to decide where to place the "doors" between each world, allowing the agents to go from one world to another. At the end of this work, a transportation management scenario is succinctly shown, illustrating how this approach could be used for traffic management.

[9] proposes a coordination model for multi-agent systems using anticipation. In this model, the agents compute the consequences of their actions using constraint networks to predict if their actions will cause "undesirable states" and avoid it. This algorithm is applied on a traffic regulation problem on the intersection scale. Vehicles are agents able to Go or to Stop, and have to cross an intersection. Gridlocks may happen in the intersection (because of left-turning vehicles), and are considered as an undesirable state that can be avoided using this algorithm.

[22] proposes an approach to build regulation plans in multi-agent traffic control using unsupervised machine learning. On a traffic intersection scenario, vehicles can move and have to avoid collisions. A traffic authority gathers the information and performs a case-based reasoning, using past experience, to determine the solution to apply in order to solve the current case. Then a norm manager translates the solutions into norms for car agents, who apply these norms using a rule engine. A reduction of the number of norms necessary to accomplish the system goals (avoiding conflict) is also performed in order to reduce the number of norms the agents have to check.

[8] proposes signals traffic regulation policies based on the real-time position of each vehicle around the intersection. These policies are based on the number of vehicles waiting in each queue around the intersection.

[2] extends the MILP single intersection management in [12], proposing an optimisation over multiple intersections. Based on the access time of a vehicle through an intersection, the mechanism computes its desired access time through the next intersection in the trajectory. Using this mechanism, the intersection controllers have knowledge about both vehicles presented in the area, and vehicles that are about to arrive. The information is then taken into account in the MILP optimisation.

In [18], R. Junges and A.L.C. Bazzan propose a novel traffic light synchronisation problem using DCOP. In this work, intersections exchange messages using DCOP algorithms to optimise over the quality of the traffic lights plan. It shows that solution quality achieved with DCOP is of good quality, but communication overhead and computational time tends to be an issue for such system.

[29] extends [28] by introducing an individual priority level for vehicles. This work has first identified a problem of single intersection optimisation, that it is, in some cases, not effective network-wide. A novel priority level depending on the vehicles' information such as destination, delay helps addressing this problem and hence leads to a better congestion control mechanism at a larger scale. However, an extensive work on how to distribute this priority correctly is needed to guarantee better performance and to deal with fairness question.

4. DISCUSSION AND POTENTIAL CHALLENGES

MAS-based approaches presented above rely on different assumptions and environment than the classic intersection control system. Classic approaches gather information through a number of loops and sensors installed in the infrastructure. They are used to measure macroscopic information about the traffic such as traffic flow, velocity and density. In contrary, MAS-based approaches rely on accurate perception of the position, velocity and desired destinations from vehicles on the network, and therefore, can model the state of the network more precisely. This assumption is reasonable in an environment where vehicles are highly equipped and have significant communication capabilities. In a future state of transportation systems, the presence of CAVs can reasonably make this type of hypothesis because CAVs are proposed with these capacities.

MAS-based approaches can further be categorised into several types, based on their control variables. Some gather precise information from vehicles to optimise the traffic light plan [3][19][18]. In contrary, others aim to regulate intersections without using traffic lights [11][27][26]. Such systems require communicating to each vehicle approaching an intersection its desired crossing time. The signal received can be similar to an individual traffic light, where each vehicle has a small time window to cross. Therefore, when opting for such solutions, vehicles must be able to react precisely to their plan (i.e. acceleration, deceleration, cross).

Regarding the level of decentralisation, most approaches presented above are centralised [11][27][12], i.e. they require a control agent that optimises over the situation and sends a solution to the traffic light system or to the vehicles. However, decentralised systems are believed to have some advantages over centralised systems when dealing with problems that are distributed by nature such as traffic regulation problems. For example, one can list the privacy by not transmitting all the data to the central agent, a better robustness due to not having a central point of failure and the parallelisation of certain computations. Furthermore, a decentralised approach would only require Vehicle-to-Vehicle communication and thus can be adopted in rural areas with minimal infrastructure level.

In terms of optimisation problems, most of traffic regulation mechanisms measure their performance using the average delay or average travel time of vehicles. This performance indicator has been widely considered one of the best since a lower average delay leads to

lower fuel consumption and greater comfort for passengers. However, when facing a realistic setup, the systems need to be able to consider more criteria. Indeed, in an urban traffic area, there are various types of vehicles, each with its own intention and each vehicle values its time differently based on their types (e.g. buses, road maintenance, other priority vehicles) or on the objectives of their travels (e.g. going to a hospital, going to work and others).

We also notice that most microscopic traffic regulation methods solve optimisation problems locally, at the scale of an intersection. These approaches can at times be greedy in certain scenarios because optimising traffic on each intersection locally does not guarantee an optimal result on a global scale. On the other hand, optimising traffic in a microscopic way over the network is highly computational expensive. Hence, a mechanism needs to be able to take into account global information, without raising additional complexity issues.

5. CONCLUSION

This paper highlighted several contributions from the multi-agent community to address traffic regulation problems on isolated intersection and on multiple intersections. We discussed in this both traffic lights optimisation approaches and microscopic approaches where each vehicle's crossing time is decided individually. These works have shown the potential of using multi-agent systems to simulate and design regulation mechanisms. We have highlighted in our discussion challenges that such models are facing in order to build a system that are close to real-life condition.

REFERENCES

[1] Abdoos, M., Mozayani, N., and Bazzan, A. L. C. Holonic multi-agent system for traffic signals control. Engineering Applications of Artificial Intelligence, 26: 1575–1587, 2013.

[2] Ashtiani, F., Fayazi, S. A., and Vahidi, A. Multi-intersection traffic management for autonomous vehicles via distributed mixed integer linear programming. In 2018 Annual American Control Conference (ACC), pages 6341–6346, 2018.

[3] Balan, G. and Luke, S. History-based traffic control. In Proceedings of the fifth international joint conference on Autonomous agents and multiagent systems, 2006.

[4] Bazzan, A. L. C., De Oliveira, D., Klügl, F., and Nagel, K. To adapt or not to adapt–consequences of adapting driver and traffic light agents. In Adaptive Agents and Multi-Agent Systems III. Adaptation and Multi-Agent Learning, pages 1–14. 2008.

[5] Camponogara, E. and Kraus Jr, W. Distributed learning agents in urban traffic control. In Progress in Artificial Intelligence, pages 324–335. 2003.

[6] Champion, A. Mécanisme de coordination multi-agent fondé sur des jeux: Application à la simulation comportementale de trafic routier en situation de carrefour. Ph.D thesis, Université de Valenciennes et du Hainaut-Cambrésis, 2003.

[7] de Campos, G. R., Falcone, P., and Sjöberg, J. Autonomous cooperative driving: a velocitybased negotiation approach for intersections crossing. In 16th International IEEE Conference on Intelligent Transportation Systems, 2013.

[8] Dinanga, E. K. and Pasin, M. Toward equitable vehicle-based intersection control in

transportation networks. 8th International Workshop on Agents in Traffic and Transportation (ATT-2014), 2014.

[9] Doniec, A., Mandiau, R., Piechowiak, S., and Espié, S. Anticipation based on constraint processing in a multi-agent context. Autonomous Agents and Multi-Agent Systems, 17:339–361, 2008.

[10] Dresner, K. and Stone, P. Sharing the road: autonomous vehicles meet human drivers. In IJCAI, 2007.

[11] Dresner, K. and Stone, P. A multiagent approach to autonomous intersection management. JAIR, 31:591–656, 2008.

[12] Fayazi, S. A., Vahidi, A., and Luckow, A. Optimal scheduling of autonomous vehicle arrivals at intelligent intersections via milp. 2017 American Control Conference (ACC), pages 4920–4925, 2017.

[13] France, J. and Ghorbani, A. A. A multiagent system for optimizing urban traffic. In Intelligent Agent Technology, 2003. IAT 2003. IEEE/WIC International Conference on, pages 411–414, 2003.

[14] Goel, A. and Kumar, P. Characterisation of nanoparticle emissions and exposure at traffic intersections through fast–response mobile and sequential measurements. Atmospheric Environment, 107:374 – 390, 2015.

[15] Grégoire, J., Bonnabel, S., and de La Fortelle, A. Optimal cooperative motionplanning for vehicles at intersections. arXiv preprint arXiv:1310.7729, 2013.

[16] Hafner, M. R., Cunningham, D., Caminiti, L., and Del Vecchio, D. Automated vehicle-to-vehicle collision avoidance at intersections. In Proceedings of World Congress on Intelligent Transport Systems, 2011.

[17] Hausknecht, M., Au, T.-C., and Stone, P. Autonomous intersection management: Multi-intersection optimization. In Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on, pages 4581–4586, 2011.

[18] Junges, R. and Bazzan, A. L. C. Evaluating the performance of DCOP algorithms in a real world, dynamic problem. In AAMAS ’08, pages 599–606, 2008.

[19] Kosonen, I. Multi-agent fuzzy signal control based on real-time simulation. Transportation Research Part C: Emerging Technologies, 11:389–403, 2003.

[20] Lee, J. and Park, B. Development and evaluation of a cooperative vehicle intersection control algorithm under the connected vehicles environment. Intelligent Transportation Systems, IEEE Transactions on, 13(1):81–90, 2012

[21] Marsa-Maestre, I., de la Hoz, E., Jonker, C. M., and Klein, M. Using graph properties to enable better negotiations in complex self-interested networks. In The International Workshop on Agent-based Complex Automated Negotiations (ACAN), 2015.

[22] Morales, J., López-Sánchez, M., and Esteva, M. Using experience to generate new regulations. In IJCAI. 2011.

[23] Schepperle, H. and Böhm, K. Agent-based traffic control using auctions. In Cooperative

Information Agents XI. 2007.

[24] Schepperle, H., Böhm, K., and Forster, S. Traffic management based on negotiations between vehicles–a feasibility demonstration using agents. In Agent-Mediated Electronic Commerce and Trading Agent Design and Analysis. 2009.

[25] Schrank, D., Eisele, B., and Lomax, T. 2019 urban mobility report. 2019

[26] Tlig, M., Buffet, O., and Simonin, O. Decentralized traffic management: A synchronization-based intersection control. In ICALT, 2014.

[27] Vasirani, M. and Ossowski, S. A market-inspired approach to reservation-based urban road traffic management. In AAMAS, 2009.

[28] Vu, H., Aknine, S., and Ramchurn, S. D. A decentralised approach to intersection traffic management. In IJCAI, 2018.

[29] Vu, H., Aknine, S., Ramchurn, S. D., and Farinelli, A. Decentralised multi-intersection congestion control for connected autonomous vehicles. In EUMAS, 2020.

[30] Yan, F., Wu, J., and Dridi, M. A scheduling model and complexity proof for autonomous vehicle sequencing problem at isolated intersections. In IEEE International Conference on Service Operations and Logistics, and Informatics (SOLI) 2014.

[31] Zou, X. and Levinson, D. Vehicle-based intersection management with intelligent agents. In ITS America Annual Meeting. Minneapolis, Minnesota, 2003.

OPTIMIZATION FOR ONLINE PEOPLE AND PARCELS SHARE-A-RIDE TAXIS WITH TIME DEPENDENT TRAVEL TIMES

Nguyen Duong Hung[1*], Nguyen Xuan Hoai[2], Pham Quang Dung[3], Nguyen Van Son[3,4]

[1]*Vietnam Banking Academy, No 12 Chua Boc Street, Hanoi, Vietnam.*

[2]*AI Academy, No 489 Hoang Quoc Viet street, Hanoi, Vietnam.*

[3]*Ha Noi University of Science and Technology, No 1 Dai Co Viet Street, Hanoi, Vietnam.*

[4]*Academy of Cryptography Techniques, No 141 Chien Thang Road, Hanoi, Vietnam.*

[*]*Corresponding author: hungnd@hvnh.edu.vn*

Abstract: The dramatically growth of urban population lead to the number of transportation requests has increased rapidly. A good transport scheduling will bring better profits to companies while satisfying people demands and reducing negative social impact such as traffic jam and pollution. An integration of passenger and parcels transportation (share-a-ride) creates attractive business opportunities because the same transportation needs can be met with fewer vehicles and drivers, and then, reduces the negative social impact. People and parcel transportation requests are unknown beforehand and come from different time points, not surprisingly, the pattern of travel time is very different for rush hours. In this paper, we consider a variant of people and parcel share-a-ride problem (PPSARP) with the time dependent. We propose an algorithm to schedule taxis that cover real-life situations. We analyze the performance of the proposed algorithm on larges instances generated based on taxi traces from San Francisco.

Keywords: vehicle routing, share a ride, stochastic travel time.

1. INTRODUCTION

An integration of passenger and parcel transportation has potentially large economical benefits. Such the integration reduces, for example, the impact of congestion and air pollution [1], [2]. Passengers are willing to share taxi with parcels if they have small dimensions, low weight and then they do not affect the space and the travel time significantly. Effective usage of empty car seats by ride-sharing may represent an important opportunity to increase occupancy rates, hence substantially increase the efficiency of urban transportation systems, potentially reducing the afore mentioned problems.

Most ride-sharing models are based on the well-known Dial-a-Ride Problem [3]. A Dial-a-Ride Problem (DARP) consists of designing vehicle routes and schedules for a number of users who request pickup and drop-off points. However, these models did not adequately cover some real-life situations. It ignored road congestion, and the information of requests may not always be known beforehand. These issues are challenging problems that interests scientific researchers in several fields. Recently, Huang et al. [4] considered a time-dependent vehicle routing problem with path flexibility (TDVRP-PF). The authors described the deterministic models and employed a Route-Path approximation method. Following this work, Rincon-

Garcia et al. [5] presented a hybrid metaheuristic algorithm to solve the time-dependent vehicle routing problem with hard time windows. In the study of [6], a time-dependent and bi-objective vehicle routing problem with time windows (TD-BO-VRPTW) was proposed, which is a new extension of classical vehicle routing problem. In the other hand, Li et al. [7], [8], described a new Shared-a-Ride Problem (SARP) in which people and parcels are served by the same taxi network. In addition, three heuristic algorithms [8], [9], [10] are proposed to solve this problem. Nguyen et al. [11] presented the SARP in online scenarios. A new algorithm for solving the problem are designed utilizing the prediction information. In their model, the travel time between a point to another is a constant based on the average speed. Nevertheless, the pattern of travel time is very different for rush hours, and hence, the travel time is affected by the traffic conditions.

To the best of our knowledge, very few researches in the literature mentions the SARP with the time dependent. As a result, in this paper, we consider the PPSARP model which the travel time between two points is not fixed. Our model alleviates the deficiencies of the model in [11] by adding the time dependent factor. The main contributions of our new transportation model are as follows:

- We extend the model in [11] by considering the PPSARP with the time dependent (TD-PPSARP) where the travel time between two points depends on the distance between these points and the time of day.
- In the dynamic setting, the requests of parcels are also dynamic/on-demand.
- We propose a new anticipatory algorithm for scheduling taxis. The algorithm is experimented on a data set of real taxis in San Francisco.

The paper is organized as follows. In section 2, the online TD-PPSARP model is given. Our new anticipatory method for solving the dynamic version of PPSARP is described in Section 3. The experiment results are given in Section 4. Section 5 concludes the paper.

2. ONLINE TD-PPSARP PROBLEM

2.1. Formulation mathematical

As stated, in [11], Nguyen et al presented the mathematical formulation for the deterministic online PPSARP. In this paper, we consider the TD-PPSARP where the travel time is affected by the traffic conditions.

Following the notations of Nguyen et al. [11], let $\tau_b, \tau_b + 1, \dots, \tau_e$ denote a sequence of time points of the planning (e.g. accurate to 1 second). There are n people requests $r_1, r_2, \dots, r_n$ and m parcel requests $r_{n+1}, r_{n+2}, \dots, r_{n+m}$. Each request has pickup and drop-off points p_i, d_i (for all $i = 1, \dots, n + m$). In order to consider the PPSARP with the time dependent, we denote the travel time from point p_i to point p_j at time point t_i by $t(p_i, p_j, \tau_i)$ where $p_i, p_j \in \wp$ and τ_i is the time of day. Therefore, it can be necessary to use noising or randomization in the travel times. Once these random variables are known, it is possible to check whether the time window and ride time constraints are satisfied.

The TD-PPSARP problem is formulated as follows.

Input

- There are K taxis $1, \dots, K$. Each taxi k starts and terminates at points s_k and t_k which refer to a physical depot, for all $k = 1, \dots, K$).
- Denote $P = \{p_1, d_1, \dots, p_{n+m}, d_{n+m}\}$, $S = \{s_1, \dots, s_K\}$, $T = \{t_1, \dots, t_K\}$.
- $\bar{d}$(p): service duration at point p, for all $p \in P$.
- $\underline{t}(p)$ and $\bar{t}(p)$: earliest and latest allowed arrival time at point p, for all $p \in P$.
- There are Q physical parkings $1, \dots, Q$, parking q has capacity $c(q)$. To ease the formulation in which each route visits each point at most once, we introduce logical points associated with each physical parking: $LP(q)$ is the set of logical points associated with parking q (each point of $LP(q)$ refers to the physical parking q).
- Denote PK the set $LP(1) \cup \dots \cup LP(Q)$, and $\wp = P \cup S \cup T \cup PK$.
- $d(p_i, p_j)$: the distance from point p_i to point p_j, for all $p_i, p_j \in \wp$.
- $t(p_i, p_j, \tau_i)$: the travel time from point p_i to point p_j at time point τ_i, for all $p_i, p_j \in \wp$, τ_i is the time of day.

$$t(p_i, p_j, \tau_i) = \frac{d(p_i, p_j)}{\gamma} \theta(\tau_i)$$

where γ is the average speed, $\theta(\tau_i)$ has the uniform distribution $U(\vartheta_l, \vartheta_b)$ $(0 \le \vartheta_l \le 1 \le \vartheta_b)$.

Variables

- $x(p)$: the successor of point p in the routes, $\forall v \in \wp \backslash T$.
- $r(p)$: the index of taxi (route) containing point p, $\forall p \in \wp$.
- $o(p)$: the order of point p in the route $r(p)$, $\forall\, p \in \wp$.
- $l(p)$: accumulated distance of the route from the start point of route $r(p)$ to p, $\forall p \in \wp$
- $ta(p), td(p)$: arrival and departure time of taxi at point p, $\forall p \in \wp$.
- $b(q_i, \tau) = 1$ if there is a taxi at parking q_i at time point τ, and $b(q_i, \tau) = 0$, $\forall q_i \in PK$, $\tau_b \le \tau \le \tau_e$ otherwise.

Constraints

- (1) $x(p) \neq p, \forall p \in \wp \backslash T$
- (2) $l(x(p)) = l(p) + d(p, p(x)), \forall p \in \wp \backslash T$
- (3) $r(p) = r(x(p)), \forall p \in \wp \backslash T$
- (4) $o(x(p)) = o(p) + 1,, \forall p \in \wp \backslash T$
- (5) $r(p_i) = r(d_i), \forall i = 1, \dots, n + m$
- (6) $o(p_i) = o(d_i)\,, \forall i = 1, \dots, n + m$
- (7) $o(p_i) = o(d_i) - 1, \forall i = 1, \dots, n$
- (8) $ta(x(p)) = td(p) + t(p, x(p), td(p)), \forall p \in \wp \backslash T$
- (9) $td(p) = ta(p) + \bar{d}(p), \forall p \in \wp$
- (10) $td(p) > ta(p), \forall p \in PK$

- (11) $\underline{t}(p) \leq ta(p) \wedge ta(p) \leq \bar{t}(p)$
- (12) $b(q_i, \tau) = ta(q_i) \leq \tau \wedge \tau \leq td(q_i),\ \forall q_i \in PK,\ \tau \in \{\tau_b, \dots, \tau_e\}$
- (13) $\sum_{q_i \in LP(q)} b(q_i, \tau) \leq c(q), \forall q = 1, \dots, Q, \tau \in \{\tau_b, \dots, \tau_e\}$

Constraint (1) specifies that the route must go from one point to another point. Constraint (2) relates the accumulated distance between two subsequent points on a route. This constraint also eliminates the existence of subtours. Constraint (3) specifies that a point and its successor must be on the same route. Constraint (4) relates the indices of two successive points on a route. Constraint (5) states that the pickup and delivery points must be con the same route and constraint (6) specifies that the delivery point of a request must be after the pickup point of that request. Constraint (7) states that the delivery point of a people request must be right-after the pickup point. It means that the people request must be delivery in a direct way without interruption. In the fashion, passengers do not feel any inconvenience (they travel like in traditional taxis). Constraints (8)-(10) relate the arrival and departure times from each point of the routes. Constraints (12)-(13) present the constraint on the capacity of parkings.

Objective

The objective of the problem is to maximize the total benefit. In our model, the passenger requests are served in a direct way as in traditional taxis (no interruption during the passenger delivery), thus there are no discounts for passengers. The total benefit is equal to the total revenue minus the fuel cost (the fuel cost is proportional to the total travel distance). Hence, the objective to be minimized is the total travel distance: $f = \sum_{k=1}^{K} l(t_k)$.

2.2. Itinerary representation

A schedule s_k of a taxi k is represented by a sequence $\langle v_0^k, r_1^k, r_2^k, \dots, r_{l_k}^k, p_k \rangle$ in which v_0^k is the current point, $v_0^k, r_1^k, r_2^k, \dots, r_{l_k}^k$ are request points (pickup or delivery) and p_k is a parking. A service plan of a taxi k is an itinerary I_k which consists of a sequence of taxi traversal points including the points of s_k and the road points between any two consecutive points in s_k. I_k also contains information about arrival time, departure time and action at each point of the sequence. In this paper, we consider the problem of large-size realistic road networks. The computation of shortest paths between two points takes time. Hence, in our algorithm, we use approximation distances (e.g, the Manhattan distance) for deciding the schedule (i.e., the sequence of request points), but the itinerary employs shortest path and real distance on the road network.

We denote:

- $l(I_k)$: the length (number of points) of the itinerary I_k.
- $p(I_k, i)$: the i^{th} point of the itinerary I_k.
- *prefix*(I_k, i): First i points of I_k.
- *suffix*(I_k, i): Last $l(I_k) - i + 1$ points of I_k.
- $d(u, v)$: the length of the shortest path from u to v in the road network.
- $t(u, v)$: the minimum travel time along the shortest path from u to v.
- *idx*(I_k, p): index of point p in itinerary I_k.
- *pickup*(r): the pickup point of request r.
- *delivery*(r): the delivery point of request r.
- *pk*(I_k): the parking of taxi k.

- For each index i in I_k: $ta(I_k, i)$ and $td(I_k, i)$ is arrival and departure time of taxi at $p(I_k, i)$.
- *itinerary*(s_k): the function that computes and returns I_k from s_k containing full information: sequence of points, arrival and departure times at each point.

When a request arrives at each time point, the system performs a decision to reschedule itineraries of taxis. The decision takes a time Δ. After Δ, the status of a taxi (e.g., location) changes. When rescheduling, there exists a minimum index j of I_k, called the scheduled point index, such that $prefix(I_k, j)$ cannot be changed. In the addition, if there are people on board, then they must be delivered before picking up other people as they do not want to share a ride.

3. ALGORITHM

In online scenarios, information on requests is revealed online during the plan execution. A control center manages status and locations of taxis, receiving information on incoming requests and making decisions. We describe in this section the algorithms for making decision of schedule of taxis. When a new request appears, we find the nearest taxi and the maximal index of the taxi's itinerary that can pick up the new request. Our algorithm find the nearest point in set of scheduled points on the route of taxi k which can pick up within the time window of the request (see [11] for more detail). This is different from what the available taxis in [9] are defined. Our algorithm lead to a diversification of the search process.

Algorithm 1 details the framework for online execution. The time horizon T contains discrete time points (e.g, 1 second). Line 2 initializes the prediction information in which P_f is the set of predicted request pickup points in the f^{th} time frame ($\forall f = 0, \dots, \bar{f}$). These points, called popular points, are taken randomly based on the information from the prediction method in Section 3 in [11]. Each day is divided into a sequence of time frames of equal length (e.g, every 15 minutes). Line 3 updates the queues Q of people and parcel requests arriving within $[T, T + \Delta T]$. Line 6 updates status of taxis. Lines 7-13 assign an appropriate taxi for each people/parcel request in Q and update the itinerary of the taxi. Whenever a request is received, the algorithm will insert it into the itinerary of a selected taxi and change this itinerary (i.e., re-order request points and exchange sub-itineraries with other itineraries). After we establish the itinerary of the taxi, we used a greedy algorithm to re-organize the remaining requests of all taxis. In the exchange improvement function, we try to find the best exchange of the subsequence between two taxis in the hope of improving the benefit. The algorithms are our preliminary results and have been published in a conference paper [11]. An example is described in Fig.1.

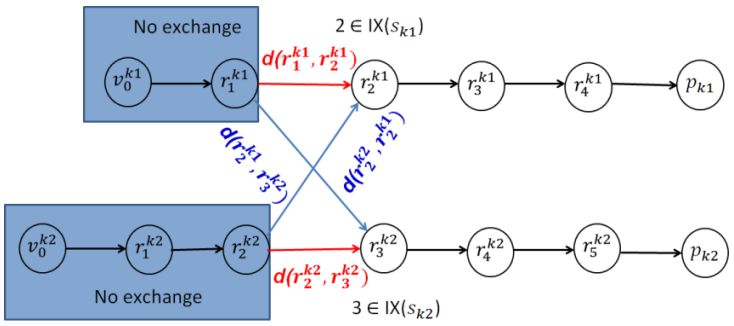


Figure 1. *The exchange improvement algorithm.*

$T \leftarrow t_b$;
Initialize P_f, $\forall f = \overline{0, \bar{f}}$;
$Q \leftarrow$ receiveRequests(T);
$T \leftarrow T + \Delta T$;
while{$T < maxSimulationTime$ **do**
    Update status of taxis;
    foreach $request\ r \in Q$ **do**
        $\langle k, j \rangle \leftarrow$ FindNearestAvailableTaxi(r; T)**;**
        if $\langle k, j \rangle \neq \perp$ **then**
            InsertRequest(r; k; j; P);
            ExchangeImprovement();
        end
    end
    $Q \leftarrow$ receiveRequests(T);
    $T \leftarrow T + \Delta T$;
end

Algorithm 1. *Framework for Online Plan Execution*

$v_0^k \leftarrow p(I_k\, j)$;
$I_k^1 \leftarrow prefix(I_k, j)$
$i \leftarrow$ minimal index such that $idx(I_k, r_i^k) > j$;
$seq \leftarrow opt(v_0^k, \langle r_i^k, \dots, r_{l_k}^k, pickup(r), delivery(r) \rangle)$;
if $seq = \perp$ **then**
    return I_k;
else
    $I_k^2 \leftarrow itinerary(seq)$;
    $p_l \leftarrow p(I_k^2, l(I_k^2))$;
    $(p, p_k) \leftarrow$ FindAPopularPointAndParking($p_l, td(p_l), P$);
    if $p \neq \perp$ **then**
        $I_k^3 \leftarrow itinerary(p_l, p, p_k)$;
    else
        $I_k^3 \leftarrow itinerary(p_l, p_k)$;
    end
    return $I_k^1 :: I_k^2 :: I_k^3$;
end

Algorithm 2. *InsertRequest(r, k, j, P).*

4. EXPERIMENTS

In our work, trace of taxis of San Francisco was used. For training the prediction model, we use the trace from 07-2005 to 07-2006. For experimenting the algorithms, we use the trace

in 03-2010. The pickup and delivery points were extracted from the taxi traces. One half of the people requests is randomly chosen and converted into parcel requests with relaxed time window. The time call of a request is specified to be the pickup time subtracting 10 minutes. The late pickup time is specified to be the pickup time plus 15 minutes. The late delivery time is specified to be the time call plus the shortest travel time from the pickup to the delivery plus 30 minutes.

4.1. Simulation design

The simulation is designed to be generic to test the efficiency of different algorithms and can be extended for other dynamic VRPs. We experiment with 2 algorithms TD-PPSARP and TD-PPSARP+prediction. TD-PPSARP will direct the taxi toward a nearby parking with the time dependent, and TD-PPSARP+prediction will score parkings based on predicted information and direct the taxi toward the predicted point and the parking of the highest score. Table 1 shows the number of all requests in our experiments.

4.2. Settings

This parameter setting is as follows, where the passenger request positions are based on the given taxi traces.

- Road network: San Francisco (from OpenStreetMap with 131245 nodes and 259792 arcs).
- 34 parkings are randomly chosen based on GoogleMaps. Parkings capacity is 39 vehicles.
- 1000 taxis: depots are at parkings.
- Start working time: 0h.
- Terminate working time: 24h.
- Speed 40km/h.
- θ_1 = 1.5.
- θ_2 = 5km.
- ϑ_l = 0.85.
- ϑ_b = 1.5.

4.3. Experimental results

TD-PPSARP and TD-PPSARP+prediction algorithms were implemented in JAVA. Test instances are solved on an Intel(R) Core(TMi7-4790 CPU @ 3.660 GHz, CPU 16 GB RAM computer. Table 2 relates the served requests. We found that at most instances, most of requests are served. We found that the number of served requests of TD-PPSARP+prediction is greater than TD-PPSARP. To illustrate the benefit of planners, Fig.2 shows the benefits of two planners with and without prediction information. It specifies that the TD-PPSARP without prediction capability often accept new requests located too far from taxi's position resulting in lower benefit.

Table 1. *The number of requests in test scripts.*

Day	#peopleReqs	#parcelReqs
1	2352	1366
2	2692	1535
3	2784	1601
4	2600	1517
5	2718	1633
6	2730	1876
7	2760	2037
8	2317	1334
9	895	299

5. CONCLUSION

In this paper, we extend the problem of scheduling people and parcel share-a-ride with the time dependent. Our model proposed by taking into account the reality that the travel time between two poins depends on the distance between these points and the time of day. An algorithm for solving the online TD-PPSARP are designed utilizing the prediction information. We evaluate our proposed algorithms on data sets extracted from taxi traces of San Francisco, combined with road network of San Francisco (from OpenStreetMap.org). Our algorithm solves the problem with a large number of requests by time constraints.

For further works, we investigate other methods as well as different approaches for solving this problem. We also analyse more deeply the successive and change cases of travel times.

Table 2. *The number of rejected requests in test scripts.*

Day	TD-PPSARP		TD-PPSARP+prediction	
	#peopleReqs	#parcelReqs	#peopleReqs	#parcelReqs
1	2341	1363	**2352**	**1366**
2	2674	1526	**2692**	**1531**
3	2772	1598	**2784**	**1599**
4	2576	1515	**2600**	1515
5	2709	1630	**2718**	1630
6	2728	1850	**2730**	**1854**
7	2760	1985	2760	**2003**
8	2314	1330	**2317**	**1333**
9	891	**285**	**893**	284

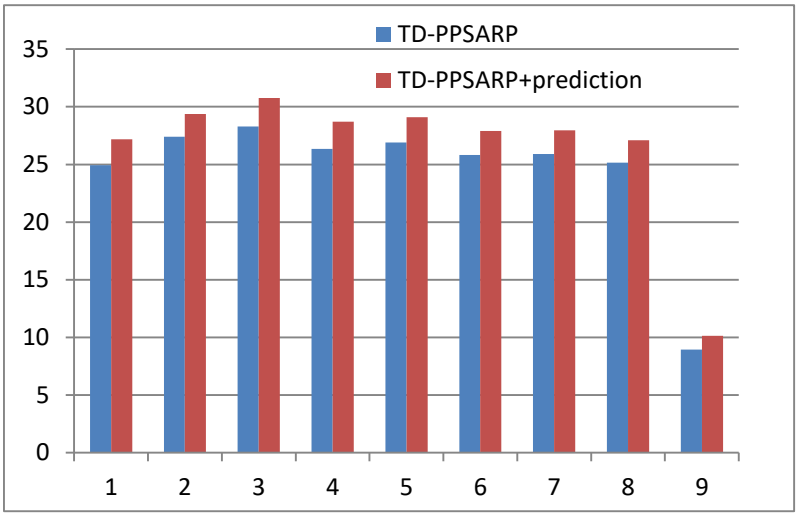


***Figure 2**. The total benefit of two planners.*

REFERENCES

[1] Lindholm, M. and Behrends, S., "Challenges in urban freight transport planning: a review in the Baltic Sea Region.," Journal of Transport Geography, p. 22, 2012.

[2] Liu, C., Kou, G., Zhou, X., Peng, Y., Sheng, H., & Alsaadi, F. E., "Time-dependent vehicle routing problem with time windows of city logistics with a congestion avoidance approach. Knowledge-Based Systems.," p. 13, 2020.

[3] Cordeau, J., & Laporte, G. T, "The dial-a-ride problem (DARP): Variants, modeling issues and algorithms.," A Quarterly Journal of Operations Research,, p. 10, 2003.

[4] Huang, Y., Zhao, L., Van Woensel, T., & Gross, J. P., "Time-dependent vehicle routing problem with path flexibility," Transportation Research Part B: Methodological, 2017.

[5] Rincon-Garcia, N., Waterson, B & Cherrett, T., "(2017). A hybrid metaheuristic for the time-dependent vehicle routing problem with hard time windows," International Journal of Industrial Engineering Computations, pp. 8(1), 141-160, 2017.

[6] Zhao, P. X., Luo, W. H., & Han, X., "Time-dependent and bi-objective vehicle routing problem with time windows," Advances in Production Engineering & Management, p. 14, 2019.

[7] Li, B., Krushinsky, D., Reijers, H. A., & Van Woensel, T, "The share-a-ride problem: People and parcels sharing taxis.," European Journal of Operational Research,, p. 12, 2014.

[8] B. Li, D. Krushinsky, W. Van, and H. Reijers., "An Adaptive Large Neighborhood Search Heuristic for the Share-a-Ride Problem. Computers & Operations Research," Computers & Operations Research, p. 12, 2016.

[9] Ngoc Quang Nguyen, Nguyen Viet Dung Nghiem, Phan Thuan DO, Khac Tuan LE, Minh Son Nguyen, Naoto MUKAI. , "People and parcels sharing a taxi for Tokyo city.," in Proceeding

SoICT 2015 Proceedings of the Sixth International Symposium on Information and Communicatio, TOKYO, 2015.

[10] V. S. NGUYEN, Q. Dung PHAM and A. T. PHAN, "An Adaptive Large Neighborhood Search Solving Large People and Parcel Share-a-Ride Problem," in *2019 6th NAFOSTED Conference on Information and Computer Science (NICS)*, Hanoi, Vietnam, 2019.

[11] Van Son, N. G. U. Y. E. N., Babaki, B., Dries, A., Dung, P. Q., & Xuan, N, "Prediction-based optimization for online People and Parcels share a ride taxis," in *In 2017 9th International Conference on Knowledge and Systems Engineering (KSE)*, Vietnam, 2017.

VISION BASED FALL DETECTION FOR THE ELDERLY PEOPLE

Vu Quang Tuyen[1], Bui Ngoc Dung[1*], Nguyen Huu Nghia[1], Chu Tran Dai[1], Viet Dung Nguyen[2]

[1]*University of Transport and Communications, No 3 Cau Giay Street, Hanoi, Vietnam.*

[2]*Hanoi University of Science and Technology, No 1 Dai Co Viet Street, Hanoi, Vietnam.*

[*]*Corresponding author: dnbui@utc.edu.vn*

Abstract: Falls are the main reason which cause injury and death in elderly people. New technology such as wearable devices and camera can help elderly people become securer at home. In this paper, we present an approach of fall detection for elderly using computer vision and deep learning approach. In each image received from camera, people were detected and estimated the pose using Single Shot MultiBox Detector and XGBoost. The key novelty of this approach is to distinguish the fall and lying action by using Long Short Term Memory network on the previous images. This approach can detect when a person falls with high accuracy. The results prove that this approach has great ability to be applied in real life.

Keywords: fall detection, human pose estimation, XGBoost, LSTM, SSD.

1. INTRODUCTION

There are between 30% and 45% of old person who is living at home fall at least once a year and from 1/2 to 1/3 of them fall more than once per year [1]. After falling, many old people cannot stand up, without any timely support system these fall can lead to serious complications and even death. Although not all of these incidents have led to physical injuries but these incidents can cause some serious psychological consequences to elderly, such as fear of falling, loss of confidence and even loss of independence. So the fall detection for elderly person become an emerge task.

Several technologies have been developed for fall detection; however, these techniques require the elderly to wear sensor devices [2, 3]. Some elders, especially those with dementia, tend to forget to wear such devices. Therefore, fall detection using camera become important technology recently [4, 5, 6]. In [7], the camera installed on the wall to monitor the presence of people in the room and the camera can detect if a person has fallen. When a fall is detected, an alarm message is sent to the cell phone of caregiver with the picture. Some technologies use multiple camera installed around the room for fall detection [8]. They use image analysis algorithm to localize people and detect when person fall. These multiple camera methods can detect person in any occlusion area of the room. However, these methods fail to distinguish the human action fall or lying in the ground.

With the recent advances in deep learning, human pose estimation has been large used in action recognition, including fall detection. Yong Du et al. [8] used a hierarchy architecture

of multiple RNN subnets: the five lowest level subnets learning features of five body parts chosen from human physical structure and each higher level subnets learning from a subgroups of previous output features until the one last RNN layer outputting the final skeleton representation before using a perceptron layer to classify the action class. Song et al. [9] explore attention mechanism in LSTM network which allows the network to focus only on parts of the input sequence that considered to be of high importance. Pengfei Zhang et al. [10] investigate the effect of different viewpoints to the accuracy of predictions and design a network to determine the best viewpoint angle for the skeleton before feeding to action recognition LSTM network. Liu et al. [11] separate skeleton data into temporal and spatial volumes and explore the use of 3d CNN which use a different convolution operation for applying to 3-dimensional data.

In this paper, we present a fall detection system using human pose estimation using deep learning approach. Our system can immediately detect and warn the caregiver when old people fall and as a consequent this system can to minimize injury of falling. The system includes a camera placing in the room and connecting to a computer. The elderly person is detected using Single Shot MultiBox Detector, after that a rectangle is drawn around the person which can build a frame rate of height and width. At the next step, pose estimation and action classification were implemented using XGBoost and SVM. Finally, to distinguish the fall and lying in the floor, stages sequence is sliced into a time-series perform action recognition using Long Short Term Memory network.

The rest of this paper is organized as follows: section 2 describes the fall detection method. Section 3 presents the experimental results and finally section 4 shows the conclusion and future works

2. METHODOLOGY

In this section, we present a method of fall detection from video. First, the people were detected from each frame using SSD. Then the coordination of human body key points in bounding box are extracted as features using DeepPose and classify states using XGBoost. Finally, the action of people is recognized using Long Shot Term Memory network.

2.1. Human detection

Input images for the fall detection system can contain both human and background. Therefore, the first step in our method is human detection and we use Single Shot Multibox Detection (SSD) for this purpose [12]. SSD is useful for object recognition in real time, it can take one single shot to detect multiple objects in the image. The key idea of the SSD is the concept of default boxes, which represent the bounding boxes based on their sizes, aspect ratios and position across the image. The purpose of the model is to decide which of the default boxes to use for a given image and then predict offsets from chosen default boxes to obtain the final prediction. SSD network uses VGG-16 as backbone network which achieved good result on high resolution image and it replaces fully-connected layer by convolution layer for extracting feature of different resolution and reduce intermediate input size. The architecture of the SSD is shown in figure 2.

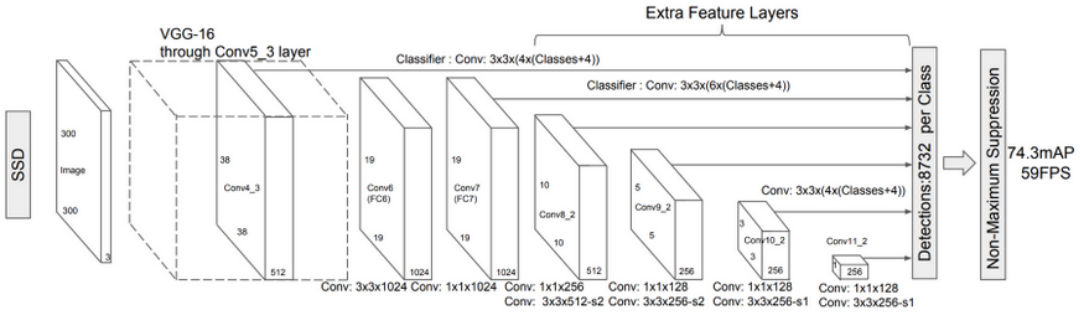


Figure 1. *Single Shot MultiBox Detector.*

2.2. Human pose estimation

Human pose estimation consider to localize the human joint by estimate coordinates for each joint in an image [13]. These connected coordinates can be used to describe the pose of the person. To express the pose, the locations of all k body joints in pose vector are encoded as y:

$$y = (\ldots, y_i^T, \ldots)^T, i \in \{1, \ldots, \mathrm{k}\}$$

For each y_i, there is x and y coordinates of ith joint. Here, Deep Pose solves pose estimation problem as a coordinate regression problem using convolution neural network (CNN) to extract feature from image and regress joints coordinate. The model consisted of an AlexNet backend including 7 layers and an extra final layer that outputs 2k joint coordinates $(x_i, y_i)*2$ for $i \in \{1, 2, \ldots, k\}$ where k is the number of joints. With the parameters θ and ψ, the model is trained using a L2 loss for regression as below:

$$\text{argmin} \sum_{(x,y)\in D} \sum_{i=1}^{k} \left\| y_i - \psi_i(x;\theta) \right\|_2^2 \text{ where D is the training set.}$$

2.3. Action recognition

Before apply action classification, sequence of poses is classified into states to further reduce search space. Human pose state classification can be solved by several methods: XGBoost, SVM, Neural Network… Input for this stage are relative position of human joints with size of x*n*1 for x is relative position of one joint. Then stages sequence is sliced into a time-series with interval length of 30 to 40 frames corresponding to 1.5 to 2.5 seconds which is suitable time for one action. Human pose after state classified can be used for action classification using Long Short Term Memory network [14].

3. RESULTS

We have tested our method using MPII dataset, including of nearly 500 different human activities, collected from YouTube videos. The activities include standing, walking, bowling, kneeling and falling. For each frame from camera, human was detected using SSD, figure 3 shows the human detected in the images. SSD can detect human in every poses faster than convolution neural network. From human in bounding box, body-model is fitted using Deep

Pose, allowing the final prediction to be human like. The results of human pose are shown in figure 4.

For the purpose of fall activity detection, the poses state outputs are limited to normal/walking state and falling state. This can be changed or expanded accordingly when applying to recognition of different actions set. To recognize the poses state, a Neural Network is employed and train on Deep Pose resulted human pose. The state recognition results are shown in figure 5.

Input is a vector contains 3D coordinates of important joints: head, left-right ears, pharynx, left-right shoulders, left-right elbows, left-right hands, body center - abdomen, hip center, left-right hip side, left-right knee, and left-right feet. As we can see in figure 3, all the join key points of person are correctly presented. When person fall down, the body key points change dramatically, and their pose are abnormal.

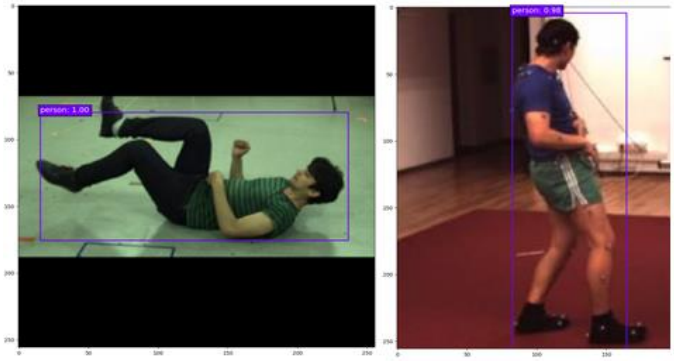


Figure 2. *Human detection using SSD.*

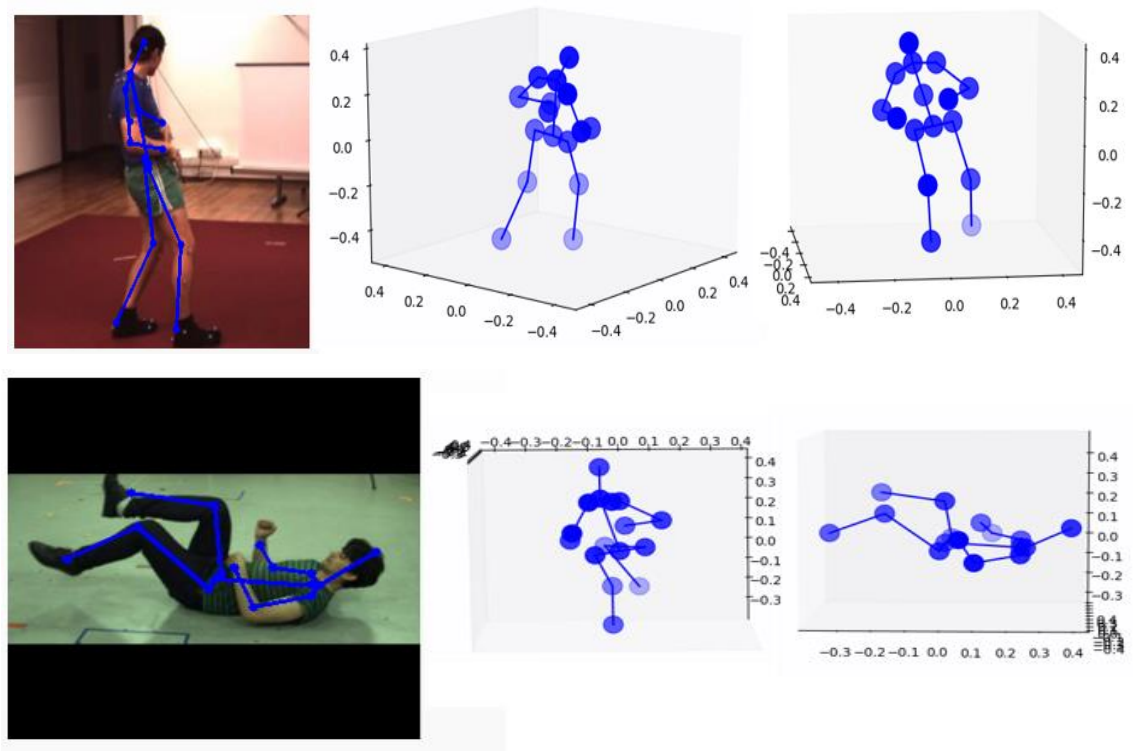


Figure 3. *Human pose estimation (up): standing state, (bottom): lying state.*

Figure 4 shows these coordinates over time axis for different pose states based on 18 points on the body. This figure shows intuitive connections that joints are often separable into two groups of upper body joints and lower body joints which reflect the pose states: walking state has more pronounce lower body activity where falling state has high coordinates changes in both joints group.

In figure 4 above, the x axis is the number of data, y axis describes the normalized relative position of each point on the body. In the walking state on the left, the position among points in the body are difference, while in the fall state, the relative position in the body is quite the same.

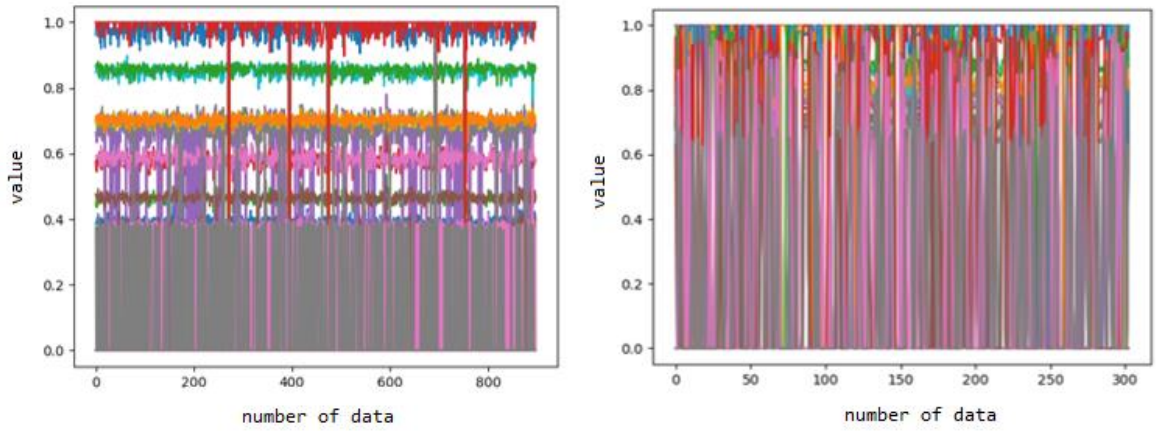


Figure 4. *Joint coordinate over time plot. (left) walking state, (right) falling state.*

After training, the poses state recognition which reached approximately 97% for training set and 94% for validation set. At the end of epochs, two values of train and test still maintain the same trajectory, validation set accuracy are lower approximately by 5%. This might show that features learned early are more general and well applied to both set while later learned features are biased toward the training set and shows sign of overfitting.

We compare our method with fall detection using CNN method. Table 1 shows the confusion matrix of our method and CNN, the results show that two methods archive more accuracy than CNN method in all normal and fall state.

Table 1. *Confusion matrix of our method (left) and CNN method (right).*

True label	Fall	145	18
	Normal	59	714
		Predict	

True label	Fall	133	30
	Normal	86	687
		Predict	

In this table, the true positive (TP), false positive (FP), false negative (FN), true negative (TN) are presented. Our method shows the accuracy 88.95% and 93.37% in fall and normal state, respectively. The CNN method shows 81.59% for fall and 88.87% for normal state.

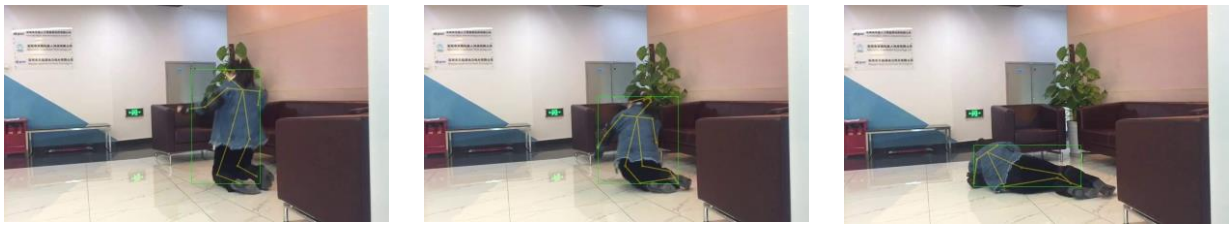


Figure 5. *Result of fall detection.*

4. CONCLUSION

This paper proposed a video-based fall detection using human pose estimation as the feature extraction method. We first used SSD and DeepPose to detect people and extract the coordination of human body key points in each frame in the video. Our method not only detect the fall state by XGBoosts, it can distinguish the fall and lying by employ the join coordinates over time and used Long Shot Term Memory to classify the fall action, that overcome the disadvantaged of CNN method. Our method can also detect multiple people falling. In the future, we may conduct more experiment of fall detection in complex environment.

REFERENCES

[1] World Health Organization. WHO Global Report on Falls Prevention in Older Age; World Health Organization:Geneva, Switzerland, 2007.

[2] Khojasteh SB, Villar JR, Chira C, González VM, de la Cal E., "Improving Fall Detection Using an On-Wrist Wearable Accelerometer", Sensor, Vol. 18, 2018.

[3] Delahoz, Yueng & Labrador, Miguel. (2014). Survey on Fall Detection and Fall Prevention Using Wearable and External Sensors. Sensors (Basel, Switzerland). 14. 19806-19842. 10.3390/s141019806.

[4] Anderson, J. Keller, M. Skubic, X. Chen, and Z. He, "Recognizing falls from silhouettes," in Engineering in Medicine and Biology Society, 2006. EMBS '06. 28th Annual International Conference of the IEEE, pp. 6388 -6391, 2006.

[5] E. Auvinet, C. Rougier, J. Meunier, A. St-Arnaud, and J. Rousseau, “Multiple cameras fall dataset,”DIRO-Universite de Montreal, Tech. Rep, vol. 1350, 2010..

[6] Auvinet, F. Multon, A. Saint-Arnaud, J. Rousseau, and J. Meunier, “Fall detection with multiple cameras: An occlusion resistant method based on 3-D silhouette vertical distribution,” Information Technology in Biomedicine, IEEE Transactions on, 2011.

[7] T. Lee and A. Mihailidis, “An intelligent emergency response system: preliminary development and testing of automated fall detection”, in Journal of Telemedicine and Telecare, vol. 11, no. 4, 2005, pp. 194 - 198.

[8] Yong Du, Wei Wang, and Liang Wang. Hierarchical recurrent neural network for skeleton based action recognition. In CVPR 2015, pages 1110–1118.

[9] S. Song, C. Lan, J. Xing, W. Zeng, and J. Liu. An End-to-End Spatio-Temporal Attention Model

for Human Action Recognition from Skeleton Data. In AAAI 2017, vol.1.

[10] Pengfei Zhang, Cuiling Lan, Junliang Xing, Wenjun Zeng, Jianru Xue, and Nanning Zheng. View Adaptive Recurrent Neural Networks for High Performance Human Action Recognition from Skeleton Data. In CVPR 2017, pages 2117–2126

[11] H. Liu, J. Tu, and M. Liu. Two-Stream 3d Convolutional Neural Network for Skeleton-Based Action Recognition. arXiv:1705.08106 [cs], May 2017.

[12] Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S. E.; Fu, C.-Y.; and Berg, A. C. 2016. SSD - Single Shot MultiBox Detector. In ECCV.

[13] S. Song, C. Lan, J. Xing, W. Zeng, and J. Liu. An End-to-End Spatio-Temporal Attention Model for Human Action Recognition from Skeleton Data. In AAAI 2017, volume 1, page 7.

[14] Grushin, Alexander & Monner, Derek & Reggia, James & Mishra, Ajay. (2013). Robust human action recognition via long short-term memory. 1-8. 10.1109/IJCNN.2013.6706797.

A METHOD FOR TRACE CLUSTERING IN PROCESS MINING

Hong-Nhung Bui[1*], Thi-Binh Khuat[1]

[1]*Banking Academy of Vietnam, No 12 Chua Boc Street, Hanoi, Vietnam.*

[*]*Corresponding author: nhungbth@hvnh.edu.vn*

Abstract: Process mining is a relatively young research discipline that is a bridge between data mining and business process modeling. The goal of process mining is to discover, monitor and improve real processes by extracting knowledge from event logs recorder by a variety of systems. In the task of process discovery, there are some problems such as generate spaghetti-like process models when the input event logs are less structured, more flexible. An approach to overcome this is to cluster the input event logs into simpler event sub-logs that can be adequately represented by a process model. Traditional approaches to trace clustering were transformed the traces into a vector space and using some distance metrics such as Euclidean distance, Jaccard distance, Generic edit distance etc. However, these approaches there is lack of order and relationship between the activities in each traces. In this paper, we propose a new trace clustering solution based on the idea of using the distance graph model for trace representation. Experimental results proved the effectiveness of the proposed solution on two measures of Fitness and Precision when compared to contemporary approaches.

Keywords: Event log, process mining, process discovering, trace clustering, distance graph model, fitness measure, precision measure.

1. INTRODUCTION

The subjects in process mining can be divided into three categories (i) discovery, (ii) conformance, and (iii) enhancement. Process discovery is the most important task which takes an event log and produces a model without using any a-priori information. An event log corresponds to a bag of process instances of a business process. The process models describe control-flow, organizational aspects, time aspects… and able to discover real processes merely based on example executions in event logs.

Today there are some algorithms for discovering process models from event logs, such as α (Wil M. P. Aalst and Boudewijn F. Dongen [1]), α+ (A.K.A de Medeiros et al. [9]), α++ (Lijie Wen et al. [17]), and other algorithms [2]. In real-life event logs may be extremely large making it difficult to handle them, and if it can be handle then the discovered process models may be diffuse and very hard to comprehend. An approach to overcome this is to cluster the traces into homogeneous sets of cases and analyzing each set of homogeneous cases separately. Each event sub-log corresponds to a sub model which is simpler to comprehend (Figure 1). Normally the approach for process model discovering includes two-phase. In the first phase, the input event log is refined, in which clustering algorithms are popularly used. In the second phase, process discovering algorithms are run on the refined event log to find out the model [6]. There exists some works following this approach [4,5,8,10,13,15,16].

The distance graph model for text processing has been proposed by Charu C. Aggarwal and Peixiang Zhao in 2013 [3]. Distance graphs of order k (k=0, 1, 2, …) for a document (a string of words) D based on the corpus C is a useful representation of D for text mining tasks [3], [7].

Because of the similar between the graph structure of process model and the Distance graph model, this work focuses on a trace clustering solution based on the idea of using the distance graph model for trace representation. This study is oriented to contribute a new solution to trace clustering1.

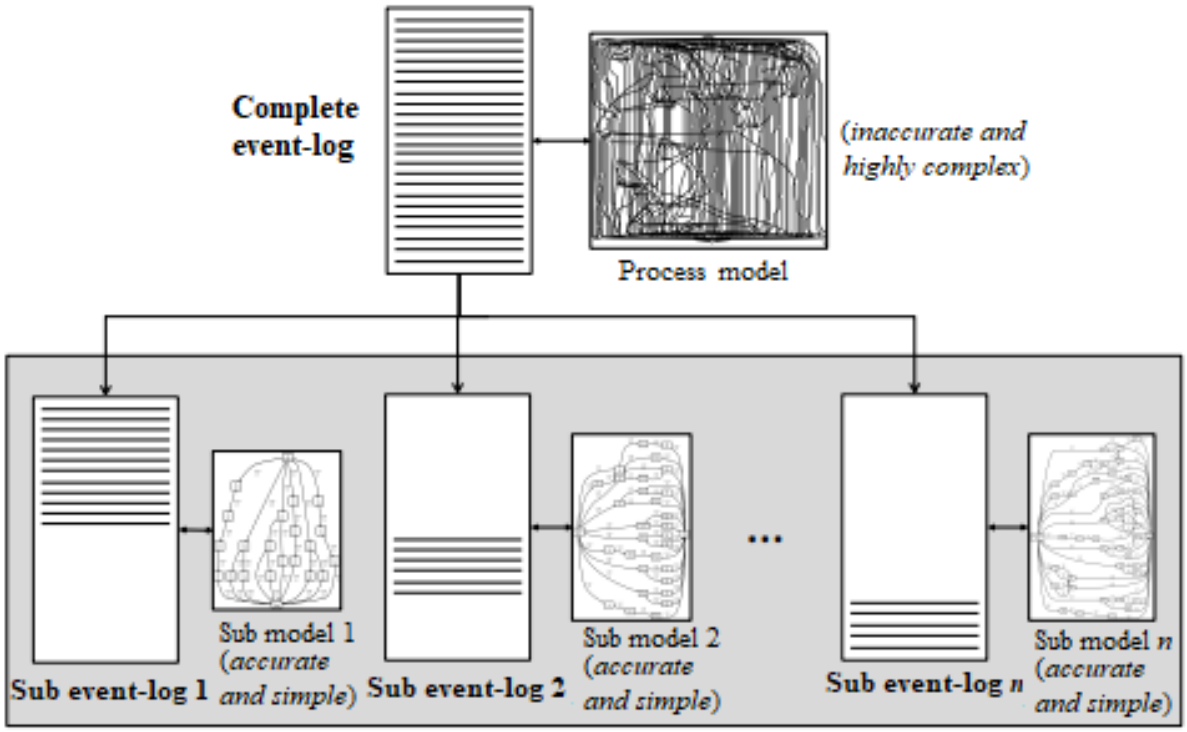


***Figure 1**. Significance of trace clustering in Process Mining.*

The rest of this article is organized as follows: In the next section, a trace clustering solution on using the distance graph model is shown. This framework includes three phases: "Trace representation and Clustering", "Process discovery", and "Model Evaluation". Experiments and remarks are described in the third section. In the fourth section, related work is introduced. And conclusions are shown in the last section.

2. APPLY THE DISTANCE GRAPH MODEL FOR TRACE CLUSTERING SOLUTION

The paper proposes a solution to trace clustering in event logs based on the distance graph model [3]. The problem is described as follows.

[1] This paper is improved from the paper "A trace clustering solution based on using the distance graph model", in proceedings of ICCCI 1 (2016) 313.

2.1 Process model discovering

Let A be the activity-name universe in an organization and A $\subseteq$ A be the set of all activity-names for a business process in the organization. A trace σ is a sequence of activities, i.e., σ ∈ A+ (where A+ is a set of non empty sequences of activities in A). Let L be a simple event log of a business process containing a set of traces constructed from A. Process discovery algorithms transform event logs into process models represented in a process modeling language, e.g. Petri nets (WorkFlow nets: WF-nets), BPMN (Business Process Modeling Notation), or YAWL (Yet Another Workflow Language), etc. There exists some algorithms for discovering process models from event logs, such as α [1], α+ [9], α++ [17], and others [2].

For example, let L = {<abdeh>, <adceg>, <acdefbdeg>, <adbeh>, <acdefdcefcdeh>, <acdeg>} (where a="register request", b="examine thoroughly", c="examine casually", d="check ticket", e="decide", f="reinitiate request", g="pay compensation", h= "reject request") be an event log for the requests for compensation business process within an airline. Figure 2 describes the WorkFlow net discovered the event log L by applying the α algorithm [2].

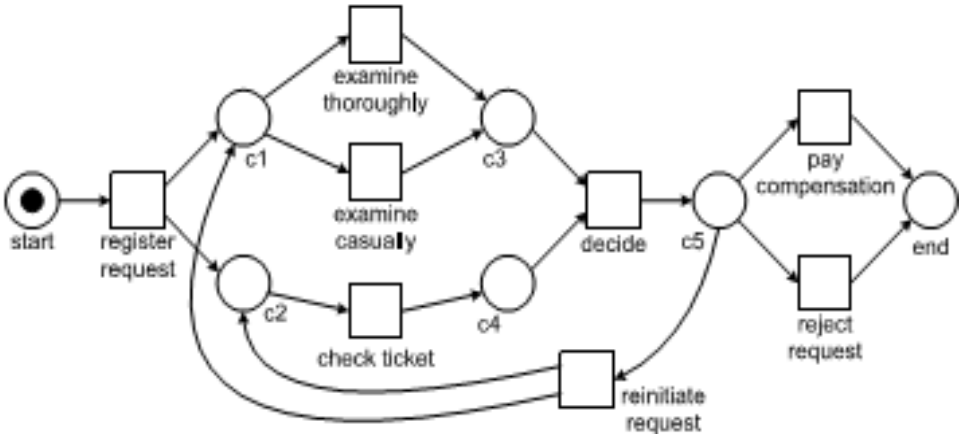


***Figure 2**. WorkFlow net discovered by the α-algorithm based on L [2].*

Normally the approach for process model discovering includes two-phase. In the first phase, the input event log is refined, in which clustering algorithms are popularly used. In the second phase, process discovering algorithms are run on the refined event log to find out the model.

Several approaches for event log refining have been proposed to transform the cases into a vector-space. The first a feature space has been defined and then each trace in the event log has been transformed into a vector using the feature space [6]. In this transformation, binary representation or frequency representation for a trace can be used. In the binary representation, we consider only the presence or absence of a feature in a trace and represent the value for that feature as 1 or 0 respectively, whereas in the frequency representation, the value for a feature corresponds to the frequency of occurrence of that feature in a trace. But in this approach there is lack of order and relationship between the activities in each traces. In some cases, the different traces also have the same representation. For example, in feature space (a,b,c,d,e,f,g,h) the traces <acdefdbeh> and <acdefbdeh> has the same binary representation (1,1,1,1,1,1,0,1) and has the same frequency representation (1,1,1,2,2,1,0,1).

2.2. The Distance Graph Model

As mentioned in the introduction section, the distance graph model ("A distance graph of order k for a document D drawn from a corpus C") for text processing was proposed by Charu C. Aggarwal and Peixiang Zhao in 2013. Figure 3 illustrates the distance graphs of orders 0, 1, and 2 for the well-known nursery rhyme "Mary had a little lamb" [3]. As stated in [3], the most common method of representing a document D is a vector of distinct terms generated from the corpus C, where each component of the vector is the frequency of a certain term appearing in D. Charu C. et al. proposed to convert a distance graph into a vector-space representation, i.e. each directed edge in the distance graph is used to create a new "token" or "pseudo-word". For example, the edge from MARRY to LITTLE (in the distance graph order 2) is used to create a new pseudo-word MARRY-LITTLE; the pseudo-word created from the edge from LAMB to itself (in the distance graph order 2) is LAMB-LAMB. The frequency of the edge is used to denote the frequency of the pseudo-word. These new pseudo-words preserve the order of words in the document, thus, when combined with distinct terms in the corpus C, they enhance the semantic of the document representation in the form of a vector.

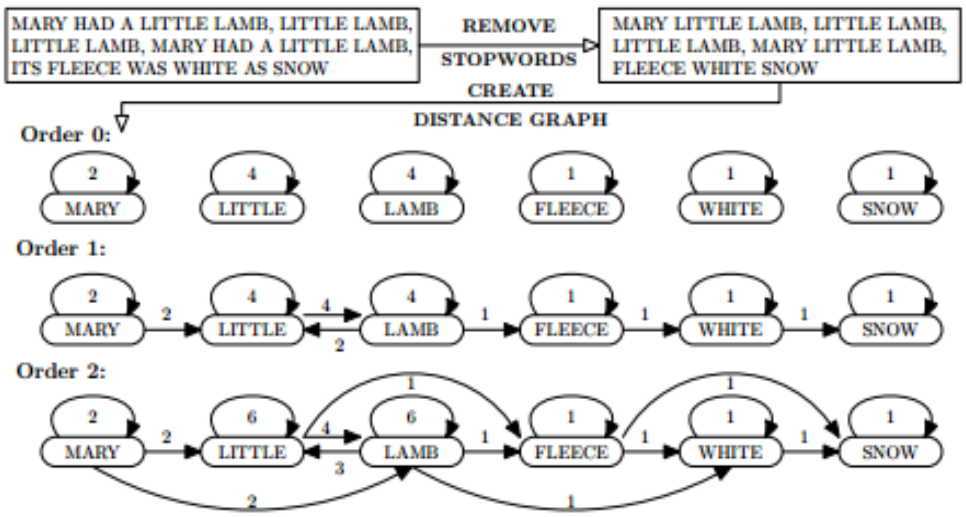


***Figure 3**. Illustration of distance graph representation [3].*

Charu C. Aggarwal and Peixiang Zhao showed some interesting features of distance graph model, as well as the effectiveness of the model applied for text classification. Since the order of activities within a trace plays an important role, one characteristic of distance graph which is considered to be suitable for trace representation is its ability to preserve the order of words in a document in the form of directed edges.

2.3. A Three-phase process discovery framework

Figure 4 describes a process discovery using trace clustering solution based on the distance graph model.

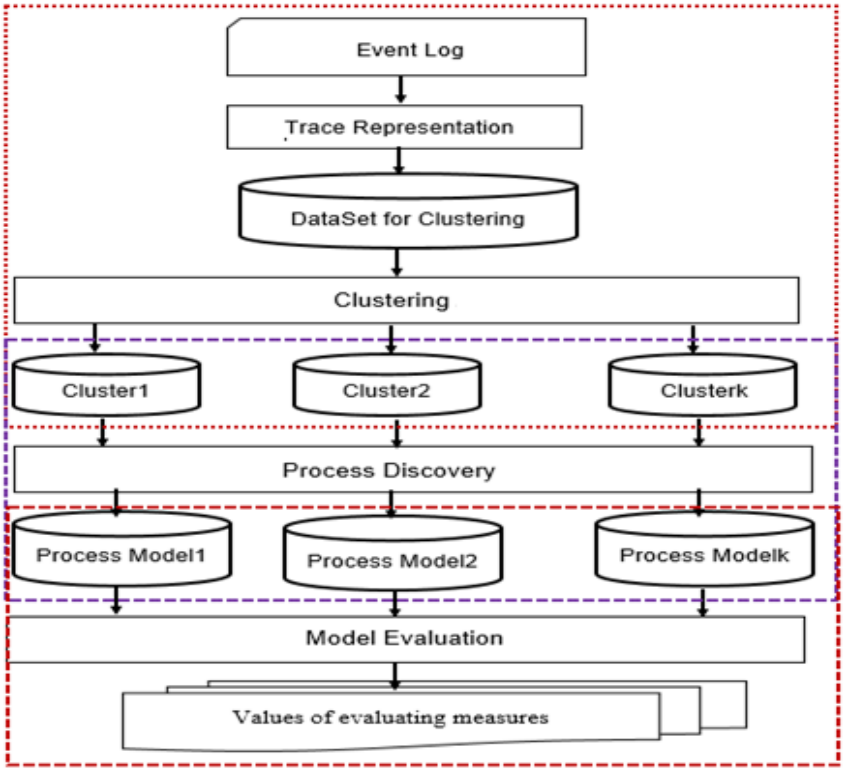


***Figure 4**. A three-phase framework of process discovery.*

The framework includes "Trace representation and Clustering", "Process discovery", and "Model evaluation" phases. Trace representation and Clustering phase includes two steps.

In the Trace Representation step, a dataset for clustering is created, in which a data point is a vector of distance graphs (with different orders) of a trace in the event log. The set A of activities in the event log is considered as the set of "distinct words" in the corpus C, and a trace in the event log is considered as a document D, thus distance graphs for a trace can be constructed.

For the given trace <acdefdbeh>:

- Order 0 distance graph is: *a*(1), *c*(1), *d*(2), *e*(2), *f*(1), *b*(1), *h*(1), where the number denotes the *frequency* of directed edges from the node to itself. This graph contains 7 unconnected components (Figure 5).

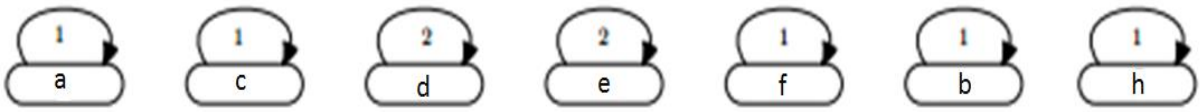


***Figure 5**. Distance graphs of order 0.*

- Order 1 distance graph is constructed from order 0 graph *a*(1), *c*(1), *d*(2), *e*(2), *f*(1), *b*(1), *h*(1) by adding more edges: *ac*(1), *cd*(1), *de*(1), *ef*(1), *fd*(1), *db*(1), *be*(1), *eh*(1), where the number denotes the frequency (Figure 6).

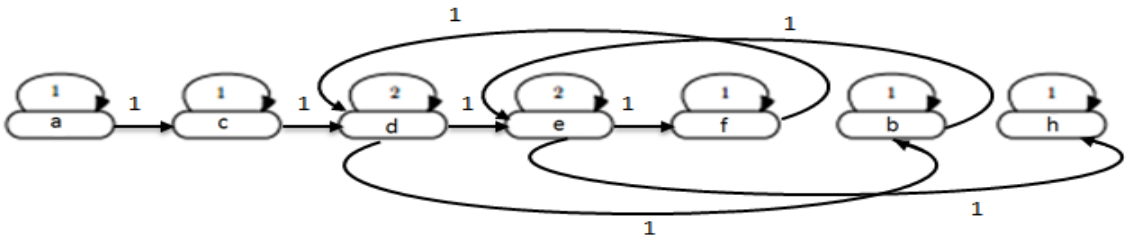


***Figure 6**. Distance graphs of order 1.*

- Order 2 distance graph is constructed from order 1 graph *a*(1), *c*(1), *d*(2), *e*(2), *f*(1), *b*(1), *h*(1), *ac*(1), *cd*(1), *de*(1), *ef*(1), *fd*(1), *db*(1), *be*(1), *eh*(1) by adding more edges: *ad*(1), *ce*(1), *df*(1), *ed*(1), *fb*(1), *de*(1), *bh*(1), where the number denotes the frequency (Figure 7).

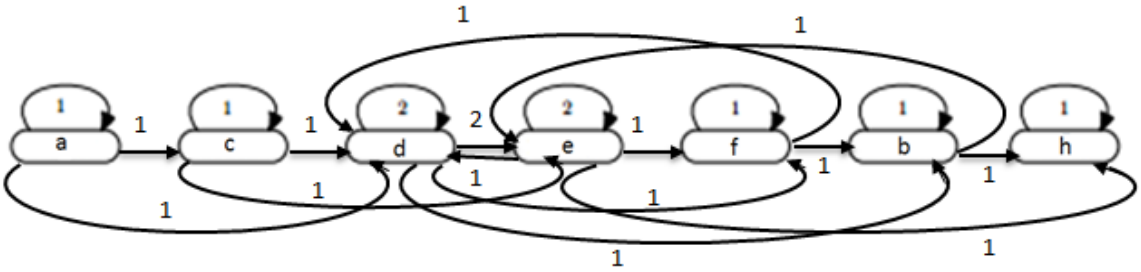


***Figure 7**. Distance graphs of order 2.*

In this trace representation the order and relationship between the activities in each traces can be deal. Moreover, the different traces will have the different representation. For example, in distance graph order of 1, the trace <acdefdbeh> has the representation (1,1,1,1,1,1,0,1,1,1, 1,1,1,1,1,1,0,0,0) and the trace <acdefbdeh> has the representation (1,1,1,1,1,1,0,1,1,1,1,1,0, 0,0,1,1,1,1).

Because event logs reflect the executions of business processes then all distance graphs of traces in an event log include some relation patterns in the discovered process model. That is why the number of features generated from all the traces in an event log L is significantly less than (|A|+|A|*(|A|-1)/2) where |A| denoted the cardinality of set A of activities.

In the Clustering step, one clustering algorithm is applied on the dataset (e.g. K-means algorithms). The output of the Trace Representation and Clustering Phase is a set of clusters (sub-logs) of traces (cases) of the event log.

In the Process Discovery Phase, a process discovery algorithm (i.e. α-algorithm) is applied on the clusters (event sub-logs) to get process models.

The Model Evaluation shows the effectiveness of result process models. Though there are four common measures for evaluation, i.e. Fitness, Precision, Generalization, and Simplicity [2], [11,12], this work considers two measures: i.e. Fitness and Precision, which had been described by A. Rozinat and Wil M. P. van der Aalst [11]. The Fitness measure indicates that the discovered model should accept the behaviors seen in the event log, and the Precision measure means that the discovered model should not accept behaviors completely unrelated to what was seen in the event log.

Since these measures are calculated on each cluster, an aggregated value for whole event log should be calculated. This work selects a weighted average value as follow:

$$w_{avg} = \sum_{1}^{k} \frac{n_i}{n} w_i \quad (1)$$

where wagv is the aggregated value of the fitness or precision measure, k is the number of clusters, n is the number of traces in the event log, ni is the number of traces in the ith cluster and wi is the value of the measure of the ith cluster.

3. EXPERIMENTAL RESULTS AND DISCUSSION

This work used the Lfull[2] and the prAm6[3] event log for experiments. This log has duplicated traces (e.g. the trace "$a\ c\ d\ e\ h$" appeared 455 times), and repeated activities in a trace (e.g. the activities d and e was repeated 2 times in the trace "$a\ c\ d\ e\ f\ d\ b\ e\ h$"). The prAm6 consists of 1200 cases with 49792 events with a small number of duplicated traces and no repeated activity. In the Trace representation step the vector-based, the distance graph order 1-based, the distance graph order 2-based, the distance graph order 3-based were used.

In the Clustering step, the K-means clustering algorithms on different number of clusters (k = 2,3,4) were used.

In Process discovery and Model evaluation phrases, ProM[4] tool was used.

3.1 The experiment with trace representation

As mentioned in the Three-phase process discovery framework section, each cases in the event log, we construct four distance graphs of order 0 (as vector-based), order 1, order 2 and order 3. Each kind of distance graphs we defines a feature space and transforms each distance graph into a binary vector in the corresponds feature space.

The experiment results are described in the Table 1. We consider the values of the dimensionality of feature space (Dim) in the cases of the vector-based, the Distance graph order 1-based, the Distance graph order 2-based, the Distance graph order 3-based trace representation.

Table 1. *Dimension of feature space in each kind of distance graph.*

	Vector	**Distance Graph order 1**	**Distance Graph order 2**	**Distance Graph order 3**
#Dim of Lfull	8	24	36	43
#Dim of prAm6	317	1076	1478	1927

3.2 The experiment with K-Means algorithm with Lfull event log

In this experiment, the K-means clustering algorithm was used to run on the vector-

[2]www.processmining.org/event_logs_and_models_used_in_book/Chapter7.zip
[3]http://data.3tu.nl/repository/uuid:44c32783-15d0-4dbd-af8a-78b97be3de49
[4]http://www.processmining.org/prom/start

based and distance graph-based trace representation. The experiment results are described in the Table 2, Table 3, Table 4 corresponding to numbers of cluster k = 2,3,4.

We also calculated the values of measures of Average-Fitness and Average-Precision (1) for activity-based (Vector) and the Distance graph-based (Distance graph) trace representation in columns titled "Avg" in the table.

***Table 2**. Experimental results of Lfull event log with number of cluster k=2.*

	Cluster 1	**Cluster 2**	**Avg**
The vector-based trace representation			
#Traces	930	461	1391
Fitness	0.982	0.995	0.987
Precision	0.805	0.987	0.817
The distance graph order 1-based trace representation			
#Traces	708	683	1391
Fitness	0.909	0.934	0.921
Precision	1	1	**1**
The distance graph order 2-based trace representation			
#Traces	961	430	1391
Fitness	0.934	1	0.955
Precision	1	0.878	0.962
The distance graph order 3-based trace representation			
#Traces	557	834	1391
Fitness	0.953	0.914	0.929
Precision	0.844	1	0.938

***Table 3**. Experimental results of Lfull event log with number of cluster k=3.*

	Cluster 1	**Cluster 2**	**Cluster 3**	**Avg**
The vector-based trace representation				
#Traces	289	461	641	1391
Fitness	0.958	1	0.999	0.991
Precision	0.626	0.840	0.750	0.754
The distance graph order 1-based trace representation				
#Traces	374	674	343	1391
Fitness	1	0.936	0.933	0.953
Precision	0.878	1	1	0.967
The distance graph order 2-based trace representation				

#Traces	128	429	834	1391
Fitness	0.923	1	0.999	**0.992**
Precision	1	1	1	**1**
The distance graph order 3-based trace representation				
#Traces	136	430	825	1391
Fitness	0.926	1	0.999	**0.992**
Precision	1	1	1	**1**

***Table 4**. Experimental results of Lfull event log with number of cluster k=4.*

	Cluster 1	**Cluster 2**	**Cluster 3**	**Cluster 4**	**Avg**
The vector-based trace representation					
#Traces	289	461	641	0	1391
Fitness	0.958	1	0.999	0	0.991
Precision	0.626	0.840	0.750	0	0.754
The distance graph order 1-based trace representation					
#Traces	535	98	277	481	1391
Fitness	0.936	0.932	0.932	0.977	0.949
Precision	1	1	1	0.871	0.956
The distance graph order 2-based trace representation					
#Traces	128	94	834	335	1391
Fitness	0.923	1	0.999	1	0.992
Precision	1	1	1	1	1
The distance graph order 3-based trace representation					
#Traces	128	193	641	429	1391
Fitness	0.923	1	0.999	1	0.992
Precision	1	1	0.750	1	0.885

The experimental results show that for most cases the performance of the distance graph based trace representation is better than that of the vector based trace representation on fitness and precision measures, especially precision measures. The results in Tables 3 showed that the suitable number of clusters for the data set is 3. The results in Tables 4 showed that in some cases, the vector based trace representation can't divide the data set into the number of clusters that we expert. It can be seen from the data in Table 3 that, on Lfull event log the Distance Graph order 2 of gives better results as compared to other trace representations on fitness and precision measures, especially on precision measure. The reason is that, the traces in Lfull event log have repeated activities and the distance graph representation can describe this relationship. On the other hand, the repeated activities in Lfull have distance 2, so the best order of distance graph is 2. The distance graph with order 3 has the same results as those of the distance graph order 2.

3.3 The experiment with K-Means algorithm with prAm6 event log

In this experiment, we experimented with a different number of clusters and found out the suitable number of clusters for the data set is 3. The results are shown in Table 5.

Table 5. *Experimental results of prAm6 event log with number of cluster k=3.*

	Cluster 1	Cluster 2	Cluster 3	Avg
The vector-based trace representation				
#Traces	275	830	95	1200
Fitness	1	0.956	0.979	**0.968**
Precision	1	0.763	0.661	**0.809**
The distance graph order 1,2,3-based trace representation				
#Traces	275	830	95	1200
Fitness	1	0.956	0.979	**0.968**
Precision	1	0.763	0.661	**0.809**

On prAm6 event log, all of the trace representations have the same clustering results. The reason is that these event logs have no repeated activity. In this case, the distance graph based trace representation is not effective.

4. RELATED WORK

G. Greco et al. [8] proposed a clustering solution on traces in event log. They used a vector representation for traces and the K-means algorithm. This work is the first study on trace clustering within the process mining domain.

R. P. Jagadeesh Chandra Bose [6], R. P. Jagadeesh Chandra Bose et al. [4,5] proposed trace clustering solutions based on using some control-flow context information. i.e. “context-aware”. The Levenshtein distance technique was used.

De Weerdt et al. [15] proposed a two phase solution to combine of trace clustering and text mining for process discovering. In the first phase, a MRA-based semi-supervised clustering technique (the SemSup-MRA algorithm) was applied. After that, there are two kinds of clusters, clusters of standard behaviors, and clusters of atypical behaviors. In the second phase, process mining and text-data mining techniques were applied. After [15], De Weerdt et al. [16] proposed the ActiTraC algorithm, a three-phase algorithm for clustering an event log into a collection of event logs (clusters). The ActiTraC algorithm includes three phases: Selection, Look ahead, and Residual trace resolution. They also developed the ActiTraCMRA algorithm, a further version of the ActiTraC algorithm.

T. Thaler et al. [14] provided a survey of trace clustering techniques. They also analyzed and compared the investigated trace clustering techniques.

This work is the first study using the distance graph model [3] for trace clustering.

5. CONCLUSIONS

In this paper, we have proposed a trace representation solution based on the distance graph model [3] to trace clustering. It was shown that the proposed clustering approach outperforms contemporary approaches to trace clustering in process mining. Compared with the previous methods, the distance graph based trace representation has better performance in case of the event log containing repeat activities. However, there are some tasks needed to do in the future. Firstly, other distance measures between graphs, e.g. distance in graph theory [18] should be studied to directly cluster traces in the form of graphs. Secondly, more clustering algorithms, especially graph-based clustering algorithms, should be considered.

REFERENCES

[1]. WMP. van der Aalst, Boudewijn F. van Dongen. Discovering Workflow Performance Models from Timed Logs. EDCIS, 45-63, 2002.

[2]. WMP. van der Aalst. Process Mining: Discovery, Conformance and Enhancement of Business Processes. Springer, 2011.

[3]. C. C. Aggarwal, Peixiang Zhao. Towards graphical models for text processing. Knowl. Inf. Syst. 36(1): 1-21, 2013.

[4]. R. P. Jagadeesh Chandra Bose, WMP. van der Aalst. Trace Clustering Based on Conserved Patterns: Towards Achieving Better Process Models. BPM Workshops 2009: 170-181, 2009.

[5]. R. P. Jagadeesh Chandra Bose, WMP. van der Aalst. Context Aware Trace Clustering: Towards Improving Process Mining Results. SDM 2009: 401-412, 2009.

[6]. R. P. Jagadeesh Chandra Bose. Process Mining in the Large: Preprocessing, Discovery, and Diagnostics. PhD Thesis, Eindhoven University of Technology, 2012.

[7]. X-Y. Dai, Chuan Cheng, Shujian Huang, Jiajun Chen. Sentiment Classification with Graph Sparsity Regularization. CICLing 2015: 140-151, 2015.

[8]. G. Greco, A. Guzzo, L. Pontieri, D. Saccà. Discovering Expressive Process Models by Clustering Log Traces. IEEE Trans. Knowl. Data Eng., 18(8): 1010–1027, 2006.

[9]. A.K.A de Medeiros, B.F. van Dongen, W.M.P. van der Aalst, A.J.M.M. Weijters. Process Mining:Extending the alpha-algorithm to Mine Short Loops. BETA Series, 2004.

[10]. A.Medeiros, A.Guzzo, G.Greco, WMP. Aalst, A.Weijters, B.Dongen, Domenico Saccà. Process Mining Based on Clustering: A Quest for Precision. BMP Workshops, 17-29, 2007.

[11]. A. Rozinat and Wil M. P. van der Aalst. Conformance checking of processes based on monitoring real behavior. Inf. Syst. 33(1): 64-95, 2008.

[12]. J. C. A. M. Buijs, Boudewijn F. van Dongen, W.M.P. van der Aalst. On the Role of Fitness, Precision, Generalization and Simplicity in Process Discovery. OTM, 305-322, 2012.

[13]. M. Song, Christian W. Günther, and Wil M. P. van der Aalst. Trace Clustering in Process Mining. Business Process Management Workshops: 109-120, 2008.

[14]. T. Thaler, Simon Felix Ternis, Peter Fettke, and Peter Loos. A Comparative Analysis of Process Instance Cluster Techniques. Wirtschaftsinformatik 2015: 423-437, 2015.

[15]. D.Weerdt, J.Broucke, S.Vanthienen, J.Baesens. Leveraging process discovery with trace

clustering and text mining for intelligent analysis of incident management processes. IEEE, 1-8, 2012.

[16]. J. De Weerdt, Seppe K. L. M. vanden Broucke, Jan Vanthienen, Bart Baesens. Active Trace Clustering for Improved Process Discovery. IEEE Trans. 25(12): 2708-2720, 2013.

[17]. L. Wen, WMP. van der Aalst, Jianmin Wang, Jiaguang Sun. Mining process models with non-free-choice constructs. Data Min. Knowl. Discov. 15(2): 145-180, 2007.

[18]. Michel M. Deza, and Elena Deza. Distances in Graph Theory. (3rd edition), Springer, 2014.

BIM APPLICATION FOR THU THIEM 2 BRIDGE IN CONSTRUCTION ENGINEERING DESIGN PHASE

Ngo Thanh Thuy[1, 2*], Do Minh Truyen[2], Nguyen Truong Giang[2], Nguyen Xuan Viet[2], Huynh Xuan Tin[1, 2]

[1] *Campus in Ho Chi Minh City, University of Transport and Communications, No. 450-451 Le Van Viet Street, Tang Nhon Phu A Ward, District 9, Ho Chi Minh City, Vietnam.*

[2]*V7 Engineering and Technology Company Limited CO.LTD, No 448/5E Le Van Viet Street, Tang Nhon Phu A Ward, District 9, Ho Chi Minh City, Vietnam.*

[*]*Corresponding author: thuynt_ph@utc.edu.vn*

Abstract: In recent years, with the development of the industrial revolution 4.0, artificial intelligence (AI) has been applied in many fields such as transportation, education, health, and security. Besides, databases (Big Data) provides electronic information applications in collecting, storing, and controlling a huge amount of information. The combination of Big Data and AI is becoming an inevitable development trend and plays an important role in modernizing the information process, creating products as well as helping to re-evaluate the implementation of the information modelling in the project, design, construction and operation phases. Thu Thiem 2 Bridge is one of forerunner project for the Building Information Modelling (BIM) applications in the design phase.

Keywords: BIM (Building Information Modeling), BIM data, Interoperability, AI (Artificial Intelligence), LOD (Level of Development).

1. INTRODUCTION

BIM is a process related to the creation and management of digital characteristics in the stages of project design, construction and operation. All individuals and organizations collaborating in the construction design can use the data in the BIM model, thereby being able to analyze prices, time and methods of construction, maintenance works. The whole process of this work is based on the sharing of information, they are always updated and supplemented continuously throughout the working process from the drafting stage to the completion of the project. With BIM, once the information is established correctly, the construction becomes faster, more accurate, and costs lower. Currently, being aware of the need for BIM in construction, most countries in the world have applied BIM in the transportation field. Vietnam is also not outside that trend, Thu Thiem 2 bridge, one of Vietnam's first implementation projects in the transport sector, is an important pilot project.

2. MATERIAL AND METHODS

2.1. Building information exchange model for the design phase

Figure 1 shows *Information exchange diagram* for BIM application of Thu Thiem 2 Bridge from the preliminary design stage to the construction engineering design phase. The

BIM model is parameterized and can be easily and quickly modified. If there is a change in design, the information of quantity, drawings, and analysis of relevant objects will be automatically updated. As a result, this process leads to save time and cost during the design phase as well as the actual construction phase.

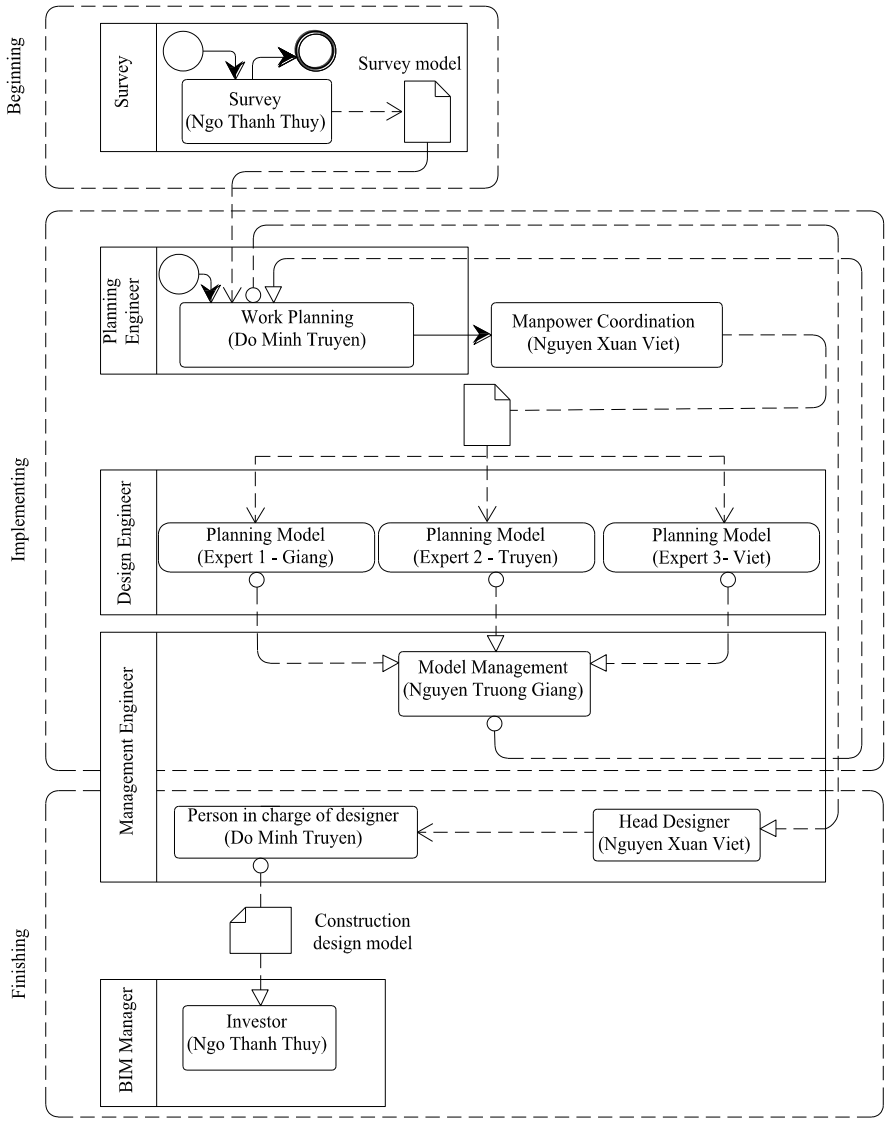


Figure 1. *Information exchange diagram in the design process [1].*

To create a BIM model, our team divides the design into three stages: Beginning, Implementation, and Finishing (Figure 1). In the first stage, the survey of the 2D plan drawings are checked and exported to a Survey Model (plan). From this Survey Model, the *Planning Engineer* will schedule the work and assign the *Manpower Coordination Engineer* to conduct a reasonable division of work; This is the implementation phase begins. Since then, a file containing the work (task) and the personnel assigned to perform the tasks will be transferred to experts. The Experts will accept their duties and make information exchange with others through *Google Drive*. When completed, the tasks will be reported to the Model Manager. If satisfactory, the manager will send information to the *Planning Engineer* to

coordinate the next tasks. This process is carried out until Construction Design Model completed. The *Planning Engineer* will send Construction Design Model to *Head Designer* for approval. If the model is met all the demands, the *Head Designer* will send it to the *Investor* [1].

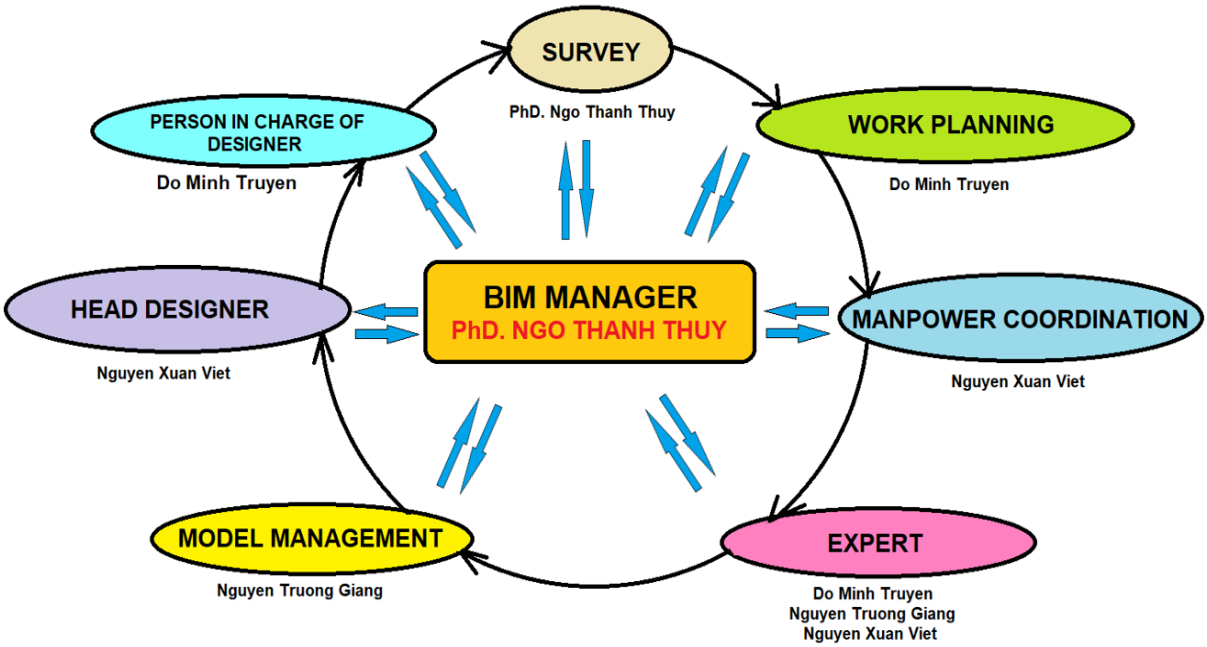


Figure 2. *Organizational chart working in the research process.*

The Thu Thiem 2 Bridge BIM application promotes cooperation in the first phase of the project among members thanks to appropriate and effective BIM information exchange, much better than traditional methods. According to Figure 2, all information is focused on one clue and everyone can catch at the same time, saving time and efficiency for tasks such as conflict checking, coordination, and adjustment design, optimize design, dissect volumes, deploy drawings, operating information, and handle complaints. Conflicting check helps reduce errors due to experience and working status of engineers. Besides, optimal design is easily carried out at all stages and rapid analysis of structure, energy, and quantity.

2.2. BIM Criteria

Modelling includes activities that involve determining who-does-what, who is responsible, and which standards need to be applied. The purpose isn't a standardized workflow, but an in-depth understanding of the data relationships between data exchange activities and stakeholders. The procedure for a specific job describes activities with specific inputs and outputs. Inputs are usually from various activities and data sources, and activities produce outputs for other subsequent activities. The output from precedent operations (e.g., preliminary design) is the input to subsequent operations (e.g., final design) of a data protocol (e.g., from software preliminary design to final design software). The information exchange scheme (Figure 1) identified in this report is based on the design-bidding-construction delivery method, and the design file has completed 100% before construction begins. Other

distribution methods such as design - construction need to be adjusted according to the information exchange scheme accordingly [2].

The FHWA-HIF-16-011 (2016) (American standards) report defines an information exchange scheme for bridge construction projects according to the design - bidding - construction process, identifying data exchange points for all stages of construction fully. This process clearly defines 5 models corresponding to 5 main stages, including the final geometry of the final element, the final structural model, the contract model, the architectural model, and the final detailed model… Information collected from TransXML, ISM1, IFC and CIS / 22 to describe the geometry, sections, materials, curved elements, etc. can be easily switched back and forth [1],[2].

The American standard has outlined the process for information exchange so that we can understand the relationships between activities to achieve the levels of data exchange, actors and data required set throughout the life of the building. The team chose a set of American standards to be applied to the Thu Thiem 2 Bridge project.

2.3. BIM Software

Modelling by Revit software follows the principles of BIM with all drawings, 2D, 3D views, and statistics, being interpretations of information from the same virtual model. When a model is built, geometric information (shape, quantity...) and non-geometry (type, code, comment, ...) is being recognized by Revit on the model.

Modelling by Revit software helps to improve the consistency and accuracy of records, especially there are adjustments in design ideas and coordination between departments. The agreement between construction works and drawings is very easy between many stakeholders (Architecture, Structure, MEP,...). Managing of the project becomes faster thanks to spending less time on a simple symbol system. Records will be easily drawn and the deployment time is extremely fast with ready-to-serve data and libraries. 3D models help to modify easily and quickly; export synchronously records; and limits errors [3].

The good communication of Revit and Lumion software will help engineers easily design more realistic perspectives of landscapes and urban contexts.

3. RESULTS

By integrating BIM technologies, conceptual designs can be represented in a digital and parameterized manner. Figure 3 shows the BIM model of the project preliminary design. This model is better understood and adopted to produce the plan, elevation and section views of bridge projects. The digitization and parameterization related to BIM models are also made the design modification more convenient. Through the research process, the team has modelled the information of Thu Thiem Bridge 2 up to LOD 350.

3.1. BIM application for detailed preliminary design (LOD 200)

Thu Thiem 2 Bridge will connect Ton Duc Thang Street – Ba Son Area – District 1 to Thu Thiem New Urban Area. The bridge includes 6 traffic lanes, the approach bridge has 4 traffic lanes and ramp bridge N1 and N2 have 1 traffic and 1 motorcycle lane. Thu Thiem 2 Main Bridge with a total length of 405m consists of the main span, back span, and side spans. The effective width of the bridge deck is 27.3m including 2x1.5 meters paved pedestrian area at edges. Model components are graphically represented in the model with relative

representations of approximate quantities, dimensions, and shapes. Non-geometric information can also be included in model components (Figure 3).

Figure 3. *Modeling Thu Thiem 2 bridge with Revit 2019 at LOD 200.*

3.2. BIM application for detailed design (LOD 350)

Model components detailed design represent the information that has been calculated and analyzed by the system of construction standards applied to the project, following the technical design phase. The information model with LOD 350 must provide enough information for material quantity, statistics, classification, arrangement, and division of construction stages (Figure 4).

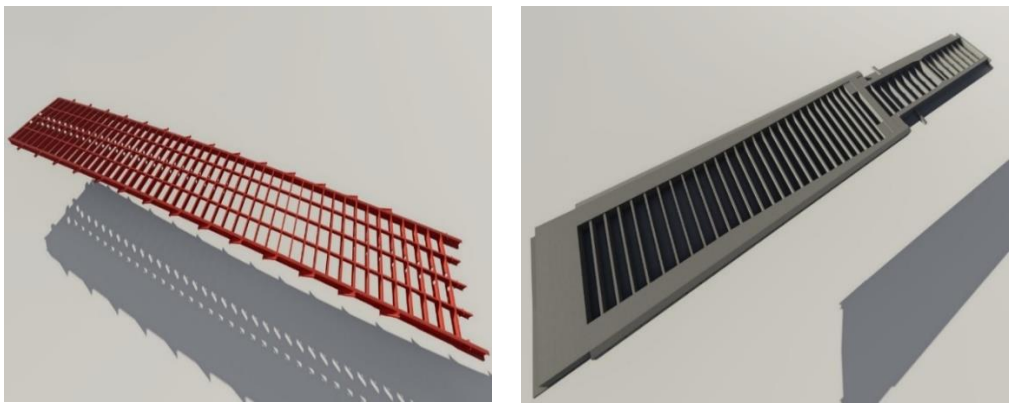


Figure 4. *Structure of steel girder span and prestressed concrete.*

Detailed design is the extension of preliminary design. The aim is to refine the integral or local structures of bridge projects, e.g., cross-sections and reinforcement detailing, and analyze the feasibility of bridge structures. By seeking to accurately portray complex structures with a detailed design, BIM can provide a useful tool for structural optimization. Using initial BIM models, 3D modeling can be deepened by combining it with detailed design drawings; this process is called detailed modeling. Figure 5 shows the detailed structural model of a bridge project. Conflict analysis and conflict reports of prestressed reinforced concrete beams are easily generated by BIM technology.

With the LOD 350, all components need the accurately coordination between the departments and related systems [4]. These components will include support details or connections. The number, size, shape, position, and direction of the components are designed and measured by the model axis without reference to notes and instructions (Figure 5).

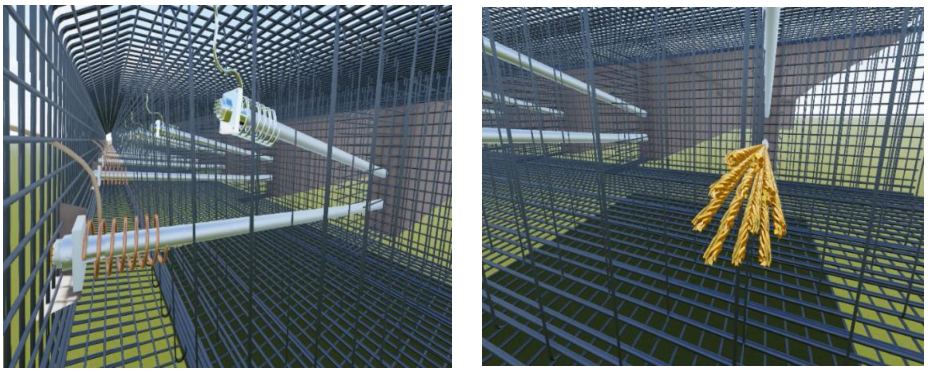
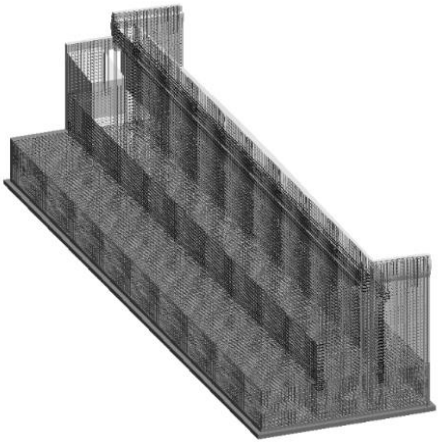


Figure 5. *The detailed model of a cable prestressed.*



A	B	C	D	E
Partition	Bar Diameter	Maximum Length	Count	Total Bar Length
B1	32 mm	7100 mm	3	355002 mm
B2	32 mm	8650 mm	1	1600264 mm
B3	32 mm	7100 mm	3	355002 mm
B4	25 mm	8660 mm	1	1602089 mm
B5	20 mm	2515 mm	4	125930 mm
B6	20 mm	6200 mm	48	6200 mm
B7	20 mm	8080 mm	16	8080 mm
B8	25 mm	2723 mm	12	122543 mm
S1	25 mm	2380 mm	4	28560 mm
S2	25 mm	1400 mm	24	1400 mm
S3	25 mm	1982 mm	68	1982 mm
S4	20 mm	3604 mm	2	43251 mm
S5	20 mm	2143 mm	36	2143 mm
T1	28 mm	6879 mm	1	1272550 mm
T2	25 mm	3957 mm	1	732132 mm
T3	16 mm	28210 mm	10	28224 mm

Figure 6. *The detailed of an abutment and exported quantity table.*

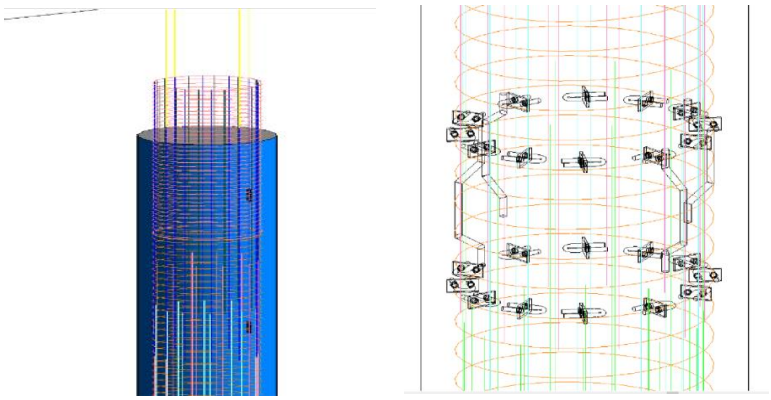


Figure 7. *The detailed of a bored cast-in-place pile D2000.*

The arrangement of reinforcement in the model (Figure 6 and Figure 7) and the assignment of specifications makes the model fully suitable for the stage of construction drawing implementation.

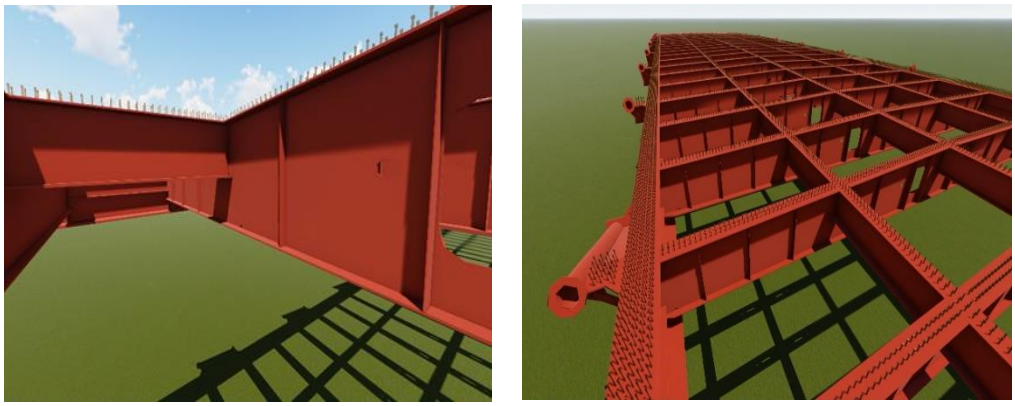


***Figure 8**. The detailed of steel girder at 350 steps.*

Steel beams are designed with detailed technical specifications and the results include the number of anchors, welding area, reinforcing ribs (Figure 8) are designed for specific limit states in Vietnamese Standard (TCVN 11823 – 2017). The model provides enough information for exporting material quantity according to construction phase (Figure 6 -7 -8).

3.3. BIM application for detailed construction design (LOD 400)

The construction model must be practical with construction measures. Through the model, quantity of materials and equipment needed for the work are high accurate exported [4]. This model shows details of construction methods with equipment and spear structures.

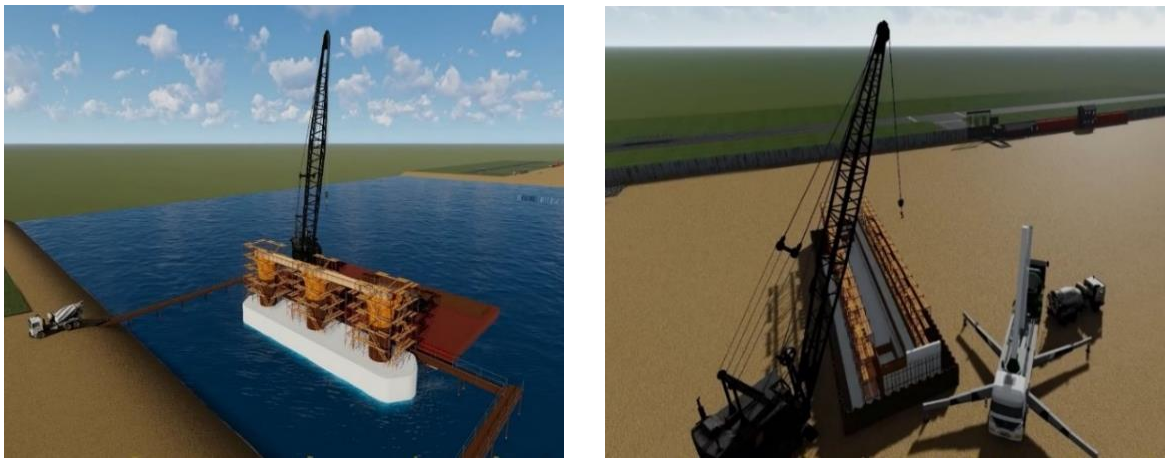


***Figure 9.** Construction model of abutment and pier.*

Construction models include high-level details and are exchanged easily with other 3D modeling software to increase reliability, safety, and eliminate conflicts. The construction model allows simulation of actual results at the site [5], [6].

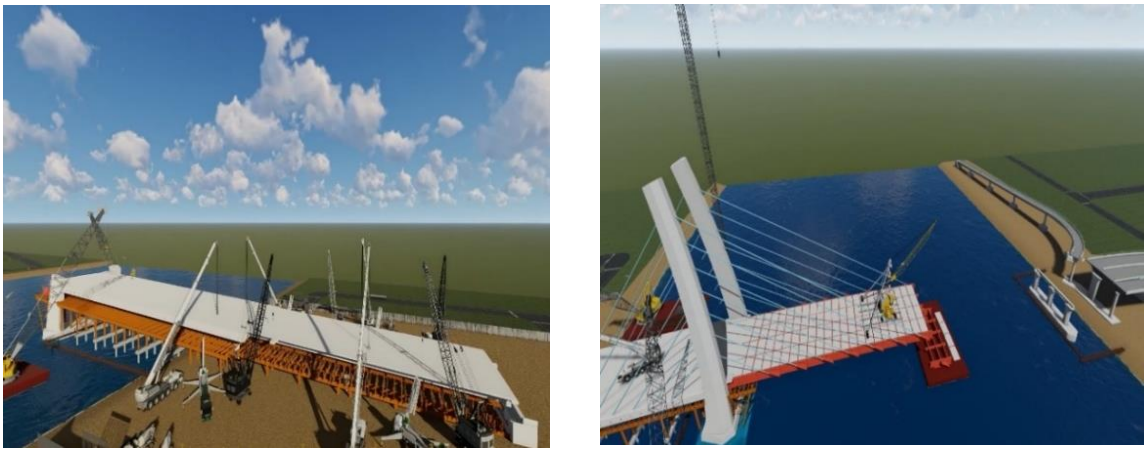


Figure 10. *Construction model of the bridge deck, tower pillars, and cable-stayed stretch.*

By using BIM and simulation technologies, the project 3D model is integrated with relevant construction methods to simulate the construction sequence of the abutment of S5 abutment, S1 pylon, S2 pylon and the entire span structure (Figure 9-10) [7],[8]. The feasibility and rationality of the construction process can be checked using construction process simulations [9]. When a construction issue is identified, the relevant construction plan is revised and simulated until an appropriate construction plan is reached. This makes the construction process smoothly and ensures the quality and progress of bridge projects [10], [11].

4. CONCLUSION

For the application of BIM in the design of bridge projects, specifically, researcher proposed an extensible information schema based on 3D parametric modelling and simulation technology to enhance the visibility of bridge design schemes. BIM application for Thu Thiem 2 Bridge in the construction engineering design phase proves that information exchange for coordination and management between the parties during project implementation becomes easy. BIM model can be used in various stages of the construction life cycle thanks to the ability of automatically generating 2D drawings, material quantity, checking conflicts, and simulation of construction execution. Not only saving time and construction costs, BIM model but also increasing construction safety. With such a BIM approach, design teams can modify in the conceptual or detailed design stage without having to start work from scratch.

ACKNOWLEDGMENT

We would like to thank the BROBIM Center, V7 Engineering & Technology CO.LTD, International conference on Applications of Artificial Intelligence in Transportation and University of Transport and Communications – Campus in Ho Chi Minh City for supporting this research.

REFERENCES

[1]. FHWA-HIF-16-011 (2016), Bridge Information Modeling Standardization Volume I: Information Exchanges, U.S Department of Transportation Federal Highway Administration.

[2]. Ths. Huỳnh Xuân Tín, Đỗ Minh Truyền, Nguyễn Xuân Việt, Nguyễn Trường Giang, Bùi Vĩ và Võ Minh Khoa (2019), "Nghiên cứu lập mô hình thông tin công trình cầu cho bước thiết kế kết cấu nhịp cầu vượt ngã tư 550 – Bình Dương", Kỷ yếu nghiên cứu khoa học sinh viên năm học 2018 – 2019, Nhà xuất bản Giao thông Vận tải, Hà Nội, tr. 3 – 7.

[3]. "BIM là gì? Ứng dụng Revit trong BIM", OneCad Việt Nam, truy cập ngày 10 tháng 02 năm 2020, < https://onecadvn.com/bim-la-gi-ung-dung-revit-trong-bim? >.

[4]. Quyết định 1057/QĐ-BXD Công bố Hướng dẫn tạm thời áp dụng Mô hình thông tin công trình (BIM) trong giai đoạn thí điểm, Bộ Xây Dựng ban hành ngày 11 tháng 10 năm 2017.

[5]. Nguyễn Việt Hùng (2015), "Nghiên cứu xây dựng lộ trình áp dụng mô hình thông tin công trình (BIM) nhằm nâng cao hiệu quả thiết kế, xây dựng và quản lý công trình tại Việt Nam", Báo cáo kết quả nghiên cứu đề tài khoa học, mã số RD 03 – 14, Bộ Xây Dựng.

[6]. "BIM được ứng dụng tại Việt Nam như thế nào" – Nguyễn Lê Minh – đăng trên website báo điện tử thuộc Bộ xây dựng ngày 12/10/2017.

[7]. Adam Strafaci (2008), "What does BIM mean for civil engineers? Road and highway projects can benefit from design using building information modeling", CE NEWS, tr. 62 – 65.

[8]. Hyun Cho, Ki Hwan Lee, Song Hyon Lee, Ta Lee, Hyun Jun Cho, Sung Hoon Kim, Sang Hyeok Nam (2011), "Introduction of Construction Management Integrated System Using BIM in the Honam High-Speed RailwayLot No. 4-2", The International Association for Automation and Robotics in Construction, 2011 Proceedings of the 28th ISARC, Seoul, Korea, tr. 1300 – 1305.

[9]. Zita Sampaio (2015), "The Introduction of the BIM Concept in Civil Engineering Curriculum", International Journal of Engineering Education, tập 31, số 1(B), tr. 305 – 305.

[10]. K. Sears, G. Sears, R. Clough (2008), "Construction Project Management: A Practical Guide to Field Construction Management (5th Edition), "John Wiley and Son, Inc., Hoboken, New Jersey.

[11]. Wenping Liu, Hongling Guo, Heng Li and Yan Li (2014), "Using BIM to Improve the Design and Construction of Bridge Projects: A Case Study of a Long-span Steel-box Arch Bridge Project", Int J Adv Robot Syst, DOI:10.5772/58442.

DETECTING SEISMIC FREQUENCY OCCUR BASED ON MOTIF DISCOVERY APPROACH

Tran Thi Dung[1*], Nguyen Thanh Son[2]

[1]*Campus in Ho Chi Minh City, University of Transport and Communications, No. 450-451 Le Van Viet Street, Tang Nhon Phu A Ward, Distric 9, Ho Chi Minh City, Vietnam*

[2]*University of Technology and Education HCMC, No 1 Vo Van Ngan Street, Linh Chieu, Thu Duc District, Ho Chi Minh City, Vietnam*

[*]*Corresponding author: dungtt_ph@utc.edu.vn*

Abstract: Detection and prediction of seismic are important problems have extensive application in many areas such as predict earthquakes, tsunamis in geography, architecture..., especially in the traffic construction domain. Researching to find out the effective seismic prediction and detection method with high accuracy is a hot trend both in Vietnam and the world. The survey phased of seismic characteristics before starting a project on traffic construction is vital. The next tasks include time series classification, frequent sequence pattern recognition, abnormal detection, and time series prediction. Motif detection in time series data has received significant recognition in the data mining community since its genesis, mainly because, motif discovery is both meaningful and more probable to succeed on big data. In this paper, the motif detection problem will be used to predict the most frequent seismic frequency. This method is being applied in many fields, in particular applied in problems with massive data volume and high efficiency. Experimental outcomes show the robustness of our method.

Keywords: time series, motif, scrimp++, time series data, seismology.

1. INTRODUCTION

Seismic is a profoundly important field of geophysics in general and earthquake science in particular. High seismic intensity is one of the main causes of damage to traffic constructions and people. Seismic research has an important role in the field of transportation construction which helps in understanding the behavior of structures of various types subjected to earthquake loads, and how we can preserve the inhabitants of that construction in an event of an earthquake. So seismic exploration is a needed task when implementing the construction of transport works. Due to the inefficiency of visually examining data recorded from devices, it is of desire to predict the seismic using data reliably [1]. Some seismic measurement methods such as refractive seismic, reflex seismic, fluorescence, However, those traditional methods become inadequate to detect seismic that occur frequency. In recent years, machine learning can be used to tackle the problem in seismic prediction.

There have been many articles on seismic and building applications such as earthquake detection [2] by using convolutional neural network, earthquake prediction [3] by using the temporal sequence of historic seismic activities in combination with the machine learning

classifiers, ... However, the problem of detecting how often the seismic frequency occurs has not been a specific study.

Time-series data mining comprises a group of intelligent techniques by which to "mine" valuable information and knowledge from time-series datasets. A time series is a collection of observations made chronologically. The nature of time-series data includes: large in data size, high dimensionality, and necessary to update continuously. Time series motifs are pairs of individual time series, or subsequences of a longer time series, which are very similar to each other.

Currently, the motif mining problems not only are being researched, developed, and deployed by famous scientists but also related to other problems in the time series.

Many problems have applied motifs discovery approaches such as understanding customers' habits, finding items with the same sales cycle, detecting copyright infringement, plagiarism detection, seismic data forecast...

Therefore, this paper investigates how much frequent seismic frequencies are possible by the application of motif detection in time series.

2. METHODOLOGY

2.1. Background

Definition 1: Time series If T is a time series then $T = (t_1, t_2, ..., t_n)$ consists of a set of n numbers with real values over time [4].

Definition 2: Sliding window: Given a time series T of length n, to determine the subsequence of length m, we use a sliding window of length m to slide through each point from left to right. on the time series T to identify each subsequence C [5].

Definition 3: Subsequence: Given a time series $T = (t_1, t_2 \ldots, t_n)$, a subqueries of length n of T is a sequence $T_{i,n} = (t_i, t_{i+1}, \ldots, t_{i+n-1})$ with $1 \leq i \leq m\text{-}n+1$ [4].

Definition 4: The motif subsequence is a pair of subqueries $\{T_{i,n}, T_{j,n}\}$ non-trivial matches of a time series most similar to T In other words, $\forall a,b,i,j$ *{*$T_{i,n}$, $T_{j,n}$*} is the subsequence motif if:*

Dist($T_{i,n}$, $T_{j,n}$*)* $\leq$ *Dist(*$T_{a,n}$,$T_{b,n}$*),* $|i-j| \geq w$ *and* $|a-b| \geq w$ *inside* $w > 0$ [4].

Note that w used in the above definition eliminates trivial matches in the case of subsequence [5] and Dist(*Ci, Cj*) is a measure of the meaningful distance between two time series.

Motif in a S time series database is a pair of different time series $\{T_i, T_j\}$, $i \neq j$, in database S has the smallest distance. Mean $\forall x, y, x \neq y, i \neq j$, DISTANCE($T_i$, T_j) $\leq$ DISTANCE(T_x, T_y) [6].

Definition 5: A Matrix distances D_i corresponding to the subsequence $T_{i,m}$ and the time series T is a vector of the Euclidean distance between a given subsequence $T_{i,m}$ and each of the time series T. Or $D_i = [d_{i,1}, d_{i,2}, ..., d_{i,n-m+1}]$, inside $d_{i,j}$ $(1 \leq j \leq n - m + 1)$ is the distance between $T_{i,m}$ and $T_{j,m}$ [7].

Definition 6: A Matrix profile P of time series T is a vector of intervals of Euclides between each subsequence of T and the nearest neighbor in T, the concept of nearest neighbor means that two pairs of subqueries have distance smallest compared to other subqueries. Or,

P = [min(D_1), min(D_2),…,min(D_{n-m+1})], inside D_i ($1 \le i \le n-m+1$) is Matrix distances D_i corresponds to the query $T_{i,m}$ and time series T [7]..

Figure 1 shows the relationship between matrix distances, Matrix distances, and Matrix profiles. Each component of the distance matrix $d_{i,j}$ is the distance between $T_{i,m}$ and $T_{j,m}$ ($1 \le i, j \le n-m+1$) in the time series T.

	D_1	D_2	…	D_{n-m+1}
D_1	$d_{1,1}$	$d_{1,2}$	…	$d_{1,n-m+1}$
D_2	$d_{2,1}$	$d_{2,2}$	…	$d_{2,n-m+1}$
…	…	…	…	…
D_{n-m+1}	$d_{n-m+1,1}$	$d_{n-m+1,2}$	…	$d_{n-m+1,n-m+1}$
	↓	↓	↓	↓
P	$min(D_1)$	$min(D_2)$	…	$min(D_{n-m+1})$

Figure 1. *Relationship between Matrix distances and Matrix profiles ([7]).*

The index i in the Matrix profile P tells us that the Euclidean distance between the subsequence $T_{i,m}$ and the nearest neighbor in the time series T. However, it does not indicate the location of the nearest neighbors, so the concept Matrix profile is given:

Definition 7: Matrix profile index I of time series T is a vector of integers: $I = [I_1, I_2, \dots I_{n-m+1}]$, where $I_i=j$ if $d_{i,j}$ = min(D_i) [7]..

Index	1	2	3	4	…	7	8	9	…	24	25	…
I	56	57	112	113	…	116	133	134	…	149	150	…

Figure 2. *Example of a Matrix profile index of a time series [7].*

The position of the minimum value in each column is stored along with the Matrix profile index.

Definition 8: 1NN-join Function is defined as the first nearest neighbor (1NN) between two subqueries *A[i]* and *B[j]*. 1NN-join functon $\theta_{1NN}(A[i], B[j])$ returns "True" if *B[j]* is the nearest neighbor of *A[i]*. 1NN-join function is a similar connection operator, applied on two sets of all subqueries; As a result, we can create *AB similarity join set*:

Definition 9: *AB* Similarity Join J_{AB} is a set of pairs of each subsequence in A with its nearest neighbor in B and vice versa.

Definition 10: Join Matrix Profile P_{ABBA} is an array of Euclidean distances for each pair in J_{ABBA}. The length of P_{ABBA} is $2 \times (n - L) + 2$ and it is twice the length of P_{AB}.

2.2. Algorithm

The motif search problem in time series data is basically divided into two branches: exact search (Exact Motif) and approximate search (Approximate Motif). Both problems have certain advantages and disadvantages. Depending on the research needs (improving efficiency or improving accuracy in finding motifs), we proceed to select the appropriate method to learn.

The SCRIMP ++ algorithm is an algorithm that combines two algorithms: PreSCRIMP and SCRIMP [7]. PreSCRIMP algorithm is an algorithm of the approximate motif search method, its complexity is $O(n^2 \log n/s)$. The SCRIMP algorithm is an algorithm of the exact search method and it is complex $O(n^2)$. The SCRIMP algorithm uses the PreSCRIMP algorithm as a time series pretreatment, it has the ability to detect motifs in the time series and it only finds an approximate Matrix Profile. From that approximate Matrix Profile will act as input for SCRIMP algorithm to find the exact Matrix Profile. That is the idea of SCRIMP ++ algorithm [7].

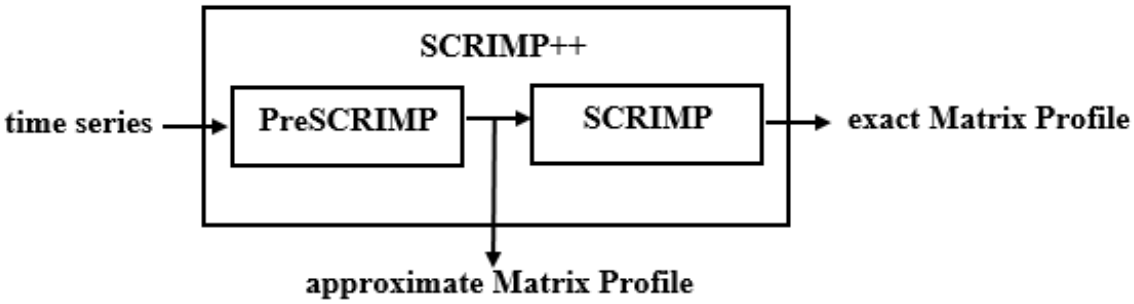


Figure 3. *SCRIMP ++ algorithm is built on two algorithms PreSCRIMP and SCRIMP [7].*

2.2.1. SCRIMP algorithm

Before going into the SCRIMP algorithm, we review the standardized formula z (the standardized formula z normalizes values in time series with amplitude) in the distance $d_{i,j}$ of the two sub-sequences $T_{i,m}$ and $T_{j,m}$ with the following formula:

$$d_{i,j} = \sqrt{2m\left(1 - \frac{Q_{i,j} - m\mu_i\mu_j}{m\sigma_i\sigma_j}\right)} \quad (2.1)$$

Inside:

+ m is the length of the subsequence
+ $Q_{i,j}$ are convolution points in $T_{i,m}$ and $T_{j,m}$
+ μ_i is the average value of $T_{i,m}$
+ μ_j is the average value of $T_{j,m}$
+ σ_i is the standard deviation of $T_{i,m}$
+ σ_j is the standard deviation of $T_{j,m}$

The input of the computation of the slippage points is a *Q* query and the time series *T*. Once done, its output will be the convolution of the points between query *Q* and all subqueries in *T*.

The process of standardizing input data is very necessary in the motif detection problem. Motif normalization helps data in time series be homogeneous during the calculation. The standardization of time series data in the SCRIMP algorithm is made easy at the beginning, and then standardized data will be included to perform the next steps. In the previously discovered motif detection algorithms, people also used data normalization before calculation, and they separated two clear steps: normalization and motif detection. Separating into two such steps will take an additional loop, which will consume more resources.

In this problem, time series T will use a sliding window of points with the length m (m is the length of the subsequence) respectively and will perform standardization of each slip. Normalizing right in the step of taking subqueries will save time, because we remove a loop to cut off subqueries, store them down and normalize each subsequence.

The SCRIMP algorithm is presented as in Table 1 below:

Table 1. *SCRIMP algorithm [7].*

	SCRIMP algorithm
	Input: *A time series T and a subsequence length m*
	Output: *Matrix profile P and matrix profile index I of time series T*
1	$n \leftarrow$ *time series length T*
2	*Calculate* μ, σ *of the time series T with the subsequence length m*
3	*Initialize initial values:* $P \leftarrow$ *infs,* $I \leftarrow$ *ones*
4	*Orders* $\leftarrow$ *RandPerm(m/4+1:n-m+1) // evaluate the random order value*
5	***for*** *k* ***in*** *Orders*
6	***for*** $i \leftarrow 1$ ***to*** *n-m+2-k*
7	***if*** $i=1$ ***do*** $q \leftarrow DotProduct(T_{1,m}, T_{k,m})$
8	***else*** $q \leftarrow q-t_{i-1}t_{i+k-2} + t_{i+m-1}t_{i+k+m-2}$
9	***end if***
10	$d \leftarrow CalculateDistance(q, \mu_i, \sigma_i, \mu_{i+k-1}, \sigma_{i+k-1})$ *(formula 2.1)*
11	***if*** $d<P_i$ *do* $P_i \leftarrow d$, $I_i \leftarrow i+k-1$ ***end if***
12	***if*** $d<P_{i+k-1}$ *do* $P_{i+k-1} \leftarrow d$, $I_{i+k-1} \leftarrow i$ ***end if***
13	***end for***
14	***end for***
15	***return P,I***

The values of distances $d_{1,k}$, $d_{2,k}$, …,$d_{n-m+2-k,n-m+1}$ are calculated in pairs; If $d_{i,i+k-1}$ (denoted by *d* in line 10) is smaller than P_i (line 11) or P_{i+k-1} (line 12), the *P* and *I* values will be updated accordingly. At any time, we can stop the algorithm to check the current *P* and *I* values, *P* and *I* can now find the motif to look for, which is appropriate in problems with data. Data is huge and takes a long time to execute.

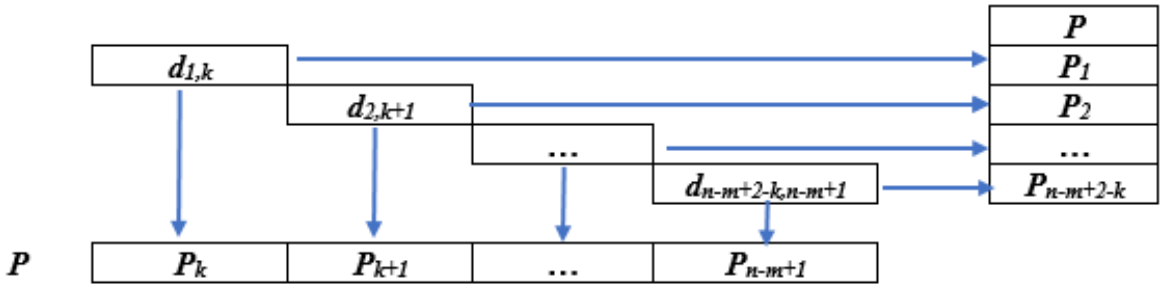


Figure 4. *Diagonal evaluation in SCRIMP algorithm[7].*

Comment:

• The motif is the exact motif

• Must wait for the algorithm to completely complete to determine the correct motif so the execution time is quite long. However, we can stop at any time to check the motif.

2.2.2. PreSCRIMP algorithm

From the limitations of an accurate motif mining algorithm, many approximation motif studies have been known for their effectiveness. From there, give a solution that focuses on the minimum points rather than focusing on all the points, which is the idea of giving up early.

Figure 2 is an example of an index profile index. Index = [*1,2,3, ..., n-m + 1*] is the position of the subsequence in the time series *T*, *I* is the matrix profile index. We can easily see that the matrix index has consecutive sequences of values. Corresponding to consecutive values in the index are consecutive values in *I* and that is also the nearest neighbors to it, which is Consecutive Neighborhood Preserving (CNP). According to CNP, if *i* and *j* are neighbors, then *i + 1* is also a neighbor of *j + 1*. In Figure 5, the 11, 12, 13 and 14 subqueries correspond to their nearest neighbors, with 136, 137, 138 and 139 chains.

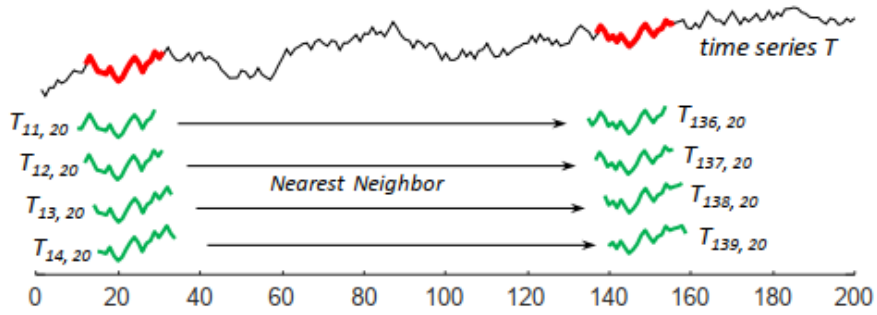


Figure 5. *Consecutive Neighborhood Preserving (CNP) attribute [7].*

Based on the CNP attribute, the idea proposed a preSCRIMP pretreatment algorithm to find an approximate matrix profile with a much faster runtime than the SCRIMP algorithm. For each sampling, we find the nearest neighbor of the sample. Suppose $T_{i,\,m}$ are a sampled substrate and $T_{j,m}$ are the nearest neighbors of $T_{i,\,m}$. According to CNP, the sequence $T_{i\,+\,k,\,m}$ can have neighbors close to it is the subsequence sequence $T_{j\,+\,k,\,m}$ *(k = -s + 1, -s + 2, ..., -2, -*

1, 1 , 2, ..., s-2, s-1) and s are sampling intervals (sampling period is equal to the length of the subsequence divided by 4).

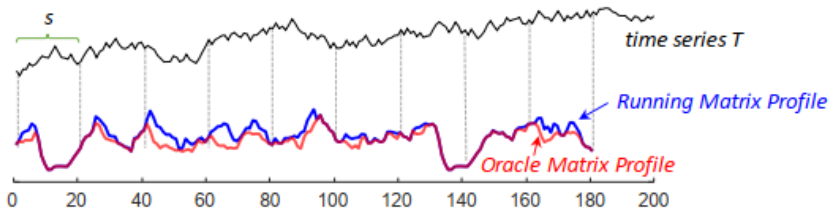


Figure 6. *Sample intervals[7].*

In the preSCRIMP algorithm [6].to find the matrix profile and matrix profile index, it uses the algorithm to calculate the distance profile which is MASS (Mueen's ultra-fast Algorithm for Similarity Search) [8]. The MASS algorithm not only returns the distance of the nearest subsequence but also the distance of all subqueries.

Table 2. *PreSCRIMP algorithm [7].*

PreSCRIMP algorithm	
Input: *A time series T, a subsequence length m and a sampling interval s* ***Output:*** *Matrix profile P and matrix profile index I of time series T*	
1	$n \leftarrow$ *time series length T,* $P \leftarrow$ *infs,* $I \leftarrow$ *ones*
2	*Calculate μ, σ of the time series T with the subsequence length m*
	for $i \leftarrow$ *RandPerm(1: s : (n-m+1))* ***do***
	$seq \leftarrow T_{i,m}$
	d=MASS (T, seq)
	P,I $\leftarrow$ *ElementWiseMin(D, P, i)*
	$P_i, I_i \leftarrow$ *min(D)*
	$j \leftarrow I_i$
	$q \leftarrow$ *CalculateDotProduct*$(P_i, \mu_i, \sigma_i, \mu_j, \sigma_j), q' = q$
	for $k \leftarrow 1$ ***to*** *min(s-1, n-m+1-max(i,j))*
	$q \leftarrow q - t_{i+k-1}t_{j+k-1} + t_{i+k+m-1}t_{j+k+m-1}$
	$d \leftarrow$ *CalculateDistance*$(q, \mu_{i+k}, \sigma_{i+k}, \mu_{j+k}, \sigma_{j+k})$ *(formula 2.1)*
	if $d < P_{i+k}$ ***do*** $P_{i+k} \leftarrow d, I_{i+k} \leftarrow j+k$ ***end if***
	if $d < P_{j+k}$ ***do*** $P_{j+k} \leftarrow d, I_{j+k} \leftarrow i+k$ ***end if***
	end for
	$q \leftarrow q'$
	for $k \leftarrow 1$ ***to*** *min(s-1, i-1, j-1)* ***do***
	$q \leftarrow q - t_{i-k+m}t_{j-k+m} + t_{i-k}t_{j-k}$
	$d \leftarrow$ *CalculateDistance*$(q, \mu_{i-k}, \sigma_{i-k}, \mu_{j-k}, \sigma_{j-k})$
	if $d < P_{i-k}$ ***do*** $P_{i-k} \leftarrow d, I_{i-k} \leftarrow j-k$ ***end if***
	if $d < P_{j-k}$ ***do*** $P_{j-k} \leftarrow d, I_{j-k} \leftarrow i-k$ ***end if***
	end for
	end for
	return P,I

After running the preSCRIMP algorithm, continue refining the matrix profile with SCRIMP algorithm, until the algorithm gives an accurate result. In the process of running SCRIMP ++ algorithm, we can stop at any time, not necessarily wait for the algorithm to run completely and the result will be an approximate Matrix profile.

2.2.3. Proposed Method

The running time of SCRIMP ++ algorithm has been significantly improved compared to the previously introduced algorithms. However, reducing the runtime to find motifs in a time series is also necessary, but we cannot stop at an algorithm. To shorten the run time, the empirical paper uses the approximation finder algorithm introduced in the **a**pproximated **m**atrix **p**rofile algorithm (AMP) [9]. The idea of AMP algorithm will be applied into SCRIMP ++ algorithm so that the execution time will be better. The APM algorithm has two main ideas: subsequence selection and number of iterations.

The improved SCRIMP ++ algorithm is implemented by two ideas:

+ Order of calculated subsequence:

The first subsequence is randomly taken to calculate the Matrix profile index. In the next loop, the nearest adjacent subsequence with the previously calculated subsequence will be selected. If the nearest adjacent v is already selected, the next random subsequence is selected. Execute the number of iterations of the second idea, then stop and return the results found.

+ Number of iterations: Based on "Birthday paradox" [10] to find out the number of loops.

The idea of the number of iterations will be based on the birthday paradox to find out, where the number of subqueries is like the number of days in the year and the motif substrate pair is like two people with the same birthday. If the two child chains have the same birthday, these two child chains are also the two children with the smallest distance from each other.

The algorithm to find the number of iterations (k):

Table 3. *Algorithm for finding loop numbers.*

Find(n,m,k)	
Input: *A time series T, subsequence length m, probability p.* ***Output:*** *Number of loops k*	
1	*k=1*
2	*prob = 0*
3	*while prob<p*
4	*prob = 1-* $e^{\frac{-k^2}{2*(n-m+1)}}$
5	*k++*
6	*end while*
7	***return k***

3. EXPERIMENTAL

Experiments in finding motifs in Seismology data are referenced on the internet published by data scientists. Experiment with two main cases: the subsequence length is 1024 points and 2048 points, the time series length is 40000 points with three algorithms to make comparisons.

Experiments show that motif position results are found in all three algorithms: SCRIMP, SCRIMP ++ and SCRIMP ++ are similar improvements. However, in time, the improved SCRIMP ++ algorithm for more optimal results, Table 4.

Loc1 and Loc2 are locations of the two motif subsequences.

Table 4. *Experimental results on seismic data.*

Len Sub	SCRIMP			SCRIMP++			SCRIMP++ improve		
	Time	Loc1	Loc2	Time	Loc1	Loc2	Time	Loc1	Loc2
1024	79.4	10252	15072	75.6	10252	15072	3.8	10252	15072
2048	60.8	10330	15150	63.3	10330	15150	2.9	10330	15150

Case 1: subsequence length is 1024 points

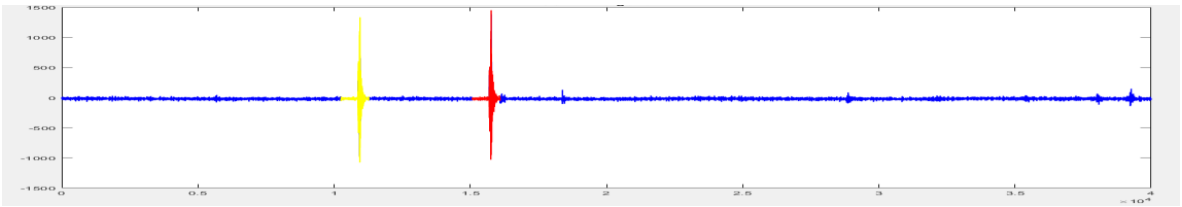


Figure 7. *Motif results with subsequence length: 1024.*

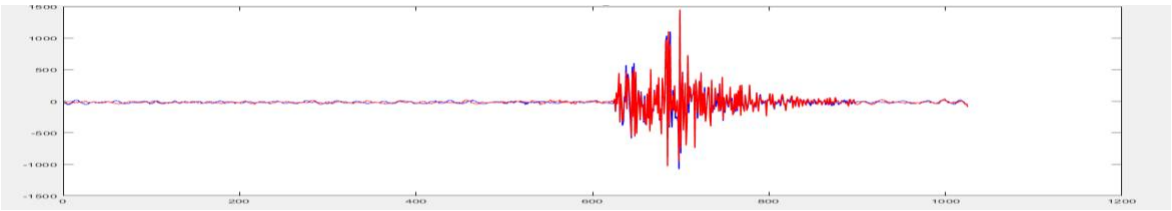


Figure 8. *Detail of Motif with subsequence length: 1024.*

Case 2: subsequence length is 2048 points

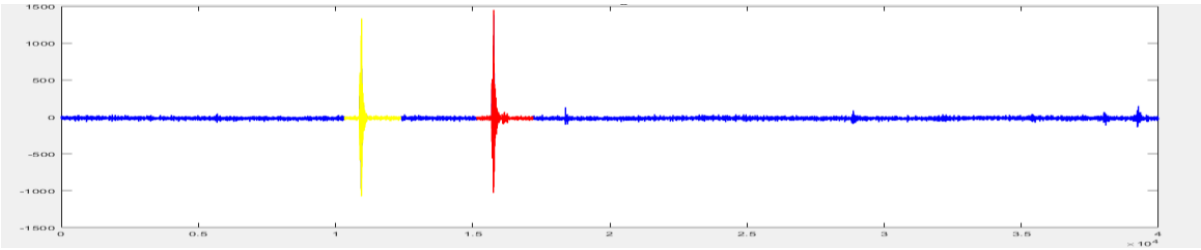


Figure 9. *Motif results with subsequence length: 2048.*

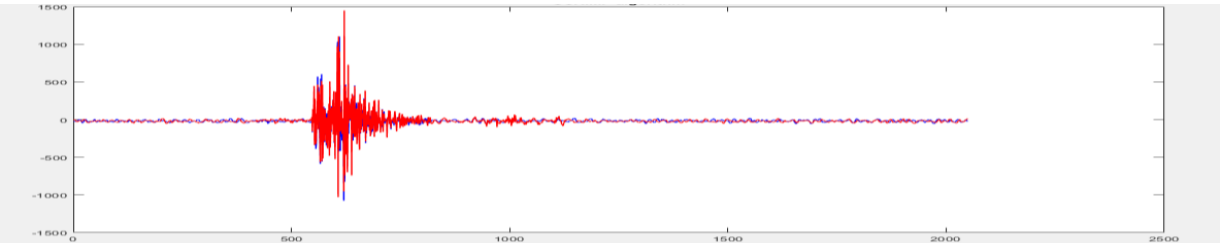


Figure 10. *Detail of Motif with subsequence length: 1024.*

Figure 7 and Figure 9 show the result of the motif sub-sequence found in seismic data. Blue is the seismic data, yellow is the first motif, and red is the second motif. Through these two figures, we can see the shape and position of the two motifs found, from which we can predict how long the common seismic frequency will be within what frequency.

In Figure 8 and Figure 10, the image comparing the difference between the two subsequences is motif, red is the first subsequence and blue is the second subsequence. If the two subsequences, motif, are found to be too different at each point, then the prediction that this is the most frequent seismic frequency may be skewed.

REFERENCES

[1] D.J. Lary, A.H. Alavi, A.H. Gandomi, A.L. Walker, Machine learning in geosciences and remote sensing, Geosci. Front., 7 (1) (2015), pp. 3-10

[2] P. Thibaut, M. Gharbi, M. Denolle, Convolutional neural network for earthquake detection and location, Science Advances, 2018.

[3] G. Alyona, G. Natalia, Machine Learning Methods for Earthquake Prediction: a Survey, SEIM, 2019.

[4] E. Keogh and S. Kasetty, On the need for time series data mining benchmark: A survey and empirical demonstration, In the 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, p. 23 – 26, 2002.

[5] E. Tufte, The visual display of quantitative information, Graphic Press, Cheshire, Connecticut.

[6] P. Salmon, J. C. Olivier, K. J. Wessels, W. Kleynhans, F. v. d. Bergh and K. C. Steenkamp, Unsupervised Land Cover Change Detection: Meaningful Sequential Time Series Analysis, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, pp. 327 335, 2011.

[7] Y. Zhu, C.C. M. Yeh, Z. Zimmerman, K. Kamgar and E. Keogh, Matrix Profile XI: SCRIMP++: Time Series Motif Discovery at Interactive, in *IEEE International Conference on Data Mining (ICDM)*, 2018.

[8] MASS: http://www.cs.unm.edu/~mueen/FastestSimilaritySearch.html

[9] K. Pariwatthanasak, C.A. Ratanamahatana, Time Series Motif Discovery Using Approximated Matrix Profile, Springer Nature Singapore, 2019.

[10] W. Ball, Other questions on probability, in *Math Recreations Essay*, 1960, p. 45

STUDY ON DESIGN AND MANUFACTURE OF INTELIGENT TRAFFIC SIGNAL CONTROL SYSTEM

Co Nhu Van[1*]

[1]*University of Transport and Communications, No 3 Cau Giay Street, Hanoi, Vietnam.*
[*]*Corresponding author: vancn@utc.edu.vn*

Abstract: The article shows the results of research, design of intelligent traffic control and monitoring system, in order to reduce the accidents and traffic congestion currently. The system applies advanced technologies and great solutions to create a complete system for practical application on the key intersection intersecting with railway at Phu Ly city, Ha Nam province[2].

Keywords: Image processing technology, artificial intelligence (AI), fuzzy control, measurement and monitoring.

1. INTRODUCTION

Generaly speaking, intelligent traffic systems and traffic signal control/monitoring systems are always a major concern for any country around the world, especially countries with high population density, residential and crowded transportation. In developed countries, the system of monitoring and control of traffic lights had been studied in early last centuries and widely applied, the system is synchronized and can monitor/control in a centralized manner throughout the domestic system. Therefore, the transport system operates effectively, minimizing traffic congestion and conflicts between vehicles participating in traffic. However, in developed countries the transport system is quite complete, plus standardized vehicles and good traffic flow as well as complete, accurate and synchronized vehicle database, due to then applying the previous algorithms to calculate the traffic flow is quite easy [3,4].

Due to the unique of traffic in Vietnam, it is difficult to apply the impoted model equipment system because our country's traffic is inequivalent and mixed among vehicles. The domestic transport infrastructure systems do not meet the standards set for foreign signal controllers. Furthermore, it is very costly for the imported system, the maintenance and repair also faces difficulties due to not mastering the technology [3,4]. Although the traffic sinal authorothies have also concentrated on their work, been working very hard to solve the traffic signal problems, but the traffic signal system is still unable to avoid errors, due to we have too many traffic lights and different types with unsynchronized. Therefore, the management and maintenance of the system also face many difficulties, especially, external factors such as storms, cable cuts, power outages, pavement, road that were constantly dug up and accidentally cut cables led to faulty or stop signaling.

***Figure 1**. Three lights are on at the sames time.*

***Figure 2**. Faulty light.*

***Figure 3**. Pictures of traffic lights are faulty, not detected early and timely processed, leading to serious traffic congestion.*

There have been a number of researches in image processing technology in Vietnam, however these researches are just providing data and information for city planners, or just stop at simulation that has not been implemented in practice. Therefore, it is very suitable to apply advanced image processing technology for the intelligent traffic light control system in Vietnam, the combination between solutions and advanced technology to create a complete control and monitoring system for traffic lights is a hot topic that is necessary to apply appropriately in Vietnam or in similar countries to bring great efficiency in economic, social, environment.

2. MAIN CONTENTS

2.1. Requirements and design the system

Requirements, function of system:

- Building a software system on the server to collect data, control, and monitor, analyze and data collection, monitor and remotely control the operation of an independent intersection or multiple intersections pre-set the light of the day or intervene from the operation center, allows the network to expand of nodes on the route, from the control center or setting special display modes in a short time at intersections such as: setting all directions are red light to reduce congestion, holding green lights in an extended direction to serve the authorities vihecles,…It also automatically check and ensure the safety of traffic lights, notify the incident center or the light signals management center, easily set the display mode when there is a fault or output signals are broken.

- Control phase diagram can be easily created for one or more continuous intersections. Unlimited number of control lines, phase design, strategy and time cycle, scalable, adjustable and quick to set up. The connecting camera system allows direct monitoring and supervision at intersections, data from the camera to the center in two ways: wired line - use for intersections near the operation monitoring center; wireless transmission - use for intersections away from the operation monitoring center, the wireless transmission applies GPRS wireless communication technology, connecting and transmitting images to the server system (operation monitoring center).

- The image processing systems will be built to measure traffic at intersections. The results of the image processing system will be sent to the control box to calculate the signal cycle in accordance with reality.

- The calculation of appropriate signal control mode based on connection of the railway crossroad warning signal, and the warning time. When the train arrives at the crossroad, the barrier closed then all directions moving into the horizontal area have a red light, while the other directions give priority to the green light [1].

- Model overview system:

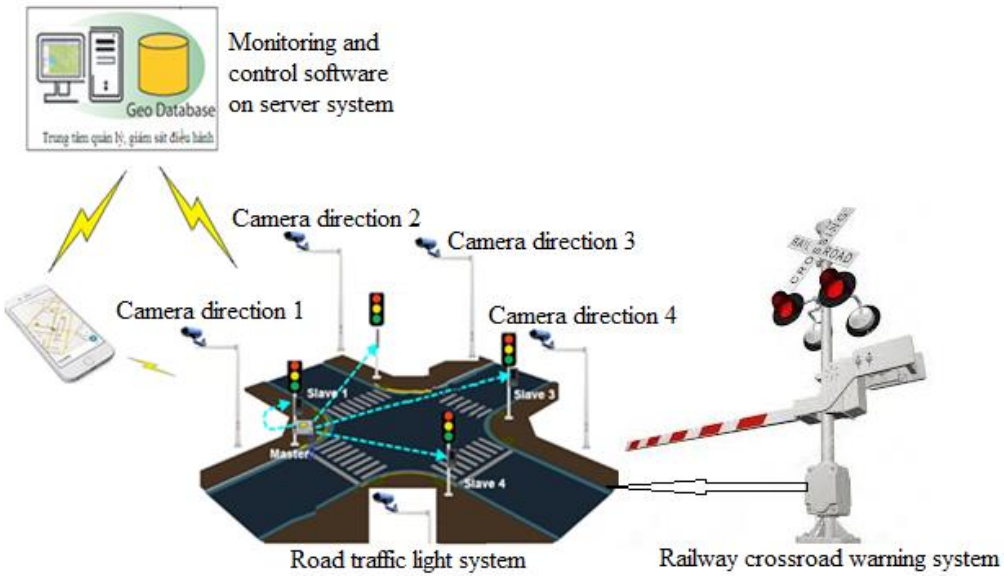


Figure 4*. System model.*

2.2. System design

2.2.1. Design of signal control box

Signal control box is the major and most important part of the system. The control box covers all functions of connecting to the center, collecting data, calculating signal cycle and controlling the operation of the system. In addition to that it also integrates the measurement unit, monitoring the operating status of the system. The overview model of the control cabinet of the system is as follows:

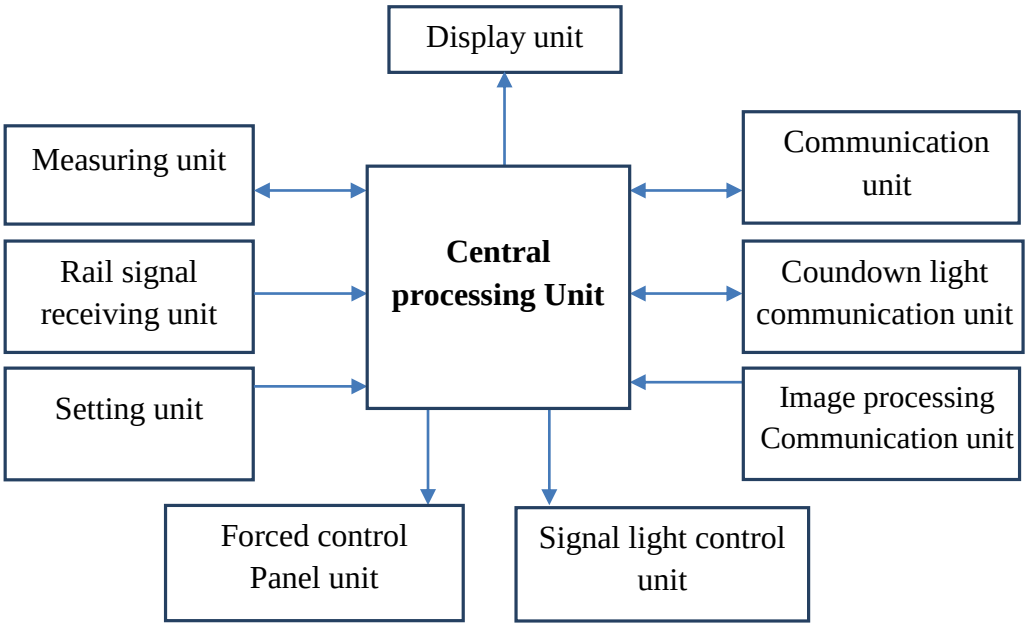


***Figure 5**. Trafficsignal control box model.*

- Central processing unit: There are full control program modes - yellow flashing warning, in order, operating on rush hour and off peak hours. All the settings of the control program parameters must allow the users to be completely active in setting via the LCD screen and button, Vietnamese language is used for interface. It allows control and monitoring of the system via SMS, GPRS, 3G networks on the basis of input devices, receiving information such as phones or websites.

- Measuring unit: Measure the current of the traffic light or cell lights at intersection to assess the operating status of the lights such as: a light that burns, collapses, or operates unstably. The power supplied to the signal light is AC current, to measure the amperage, there are many different ways - components, but the solution using shunt resistance wsr-3 0.005 is highly stable, less impact from outside and differential amplifier circuit is reliable and accurate. Monitoring the operation of the control box: power outages, power quality, can be used on/off relay or using MFM384 power meter. This is a clock that allows data transfer via RS485 standard, easy to connect to the monitoring circuit.

- Rail signal receiving unit: Tranfer as digital signal on/off, methods of controlling traffic light signals in combination with railroad, crossroads. In normal mode, the traffic light control

system operates according to the established day plans. When there is a signal from the crossroad system, the traffic light controller will switch to operate under a special plan. After the end of the crossroad signal, the system will return to normal operation according to the established day plans, it is a control program to limit vehicles entering an area where a crosswalk is blocking. Green light time from incoming directions, initially reduced slightly, after the actual run, evaluate to adjust that time, even the green light time in this direction may be zero if necessary. The method of connection and perform follow the Circular No. 28/2018/TT-BGTVT dated 14/5/2018 of the Ministry of Transport on connection of road traffic light signals with road signals at crossroads, providing information to assist the guard at the intersection of roads and railways.

- Setting unit: Seting operating values of the controller at the control box via the buttons or on the HMI screen (if using HMI). Setting values such as: light cycle of the day, yellow flashing interval, update time for monitoring information about the center, phone alert threshold (SMS).

- Forced control Panel unit: The panel is mounted right on the control panel of the signal light, which helps the operator (traffic policeman) to actively classify the lights and forced control them to prioritize routes in special cases. Operator's steps is carried out via mode switch and phase selection buttons, when turning the switch to manual control mode, all the Groups turn to the red light, press which button corresponds to that green light.

- Light signal control unit: Amplify control signal from the processor chip to 220VAC voltage to the light through triac BTA225-600B and IC buffer MOC3041. Number of control ports: 7 signal groups are expected, each group of 3 lights should have 21 ports and 4 ports to control lights.

- Image processing communication unit: Receiving the information about traffic condition through the intersections via RS232 communication.

- Coundown light communication unit: Countdown time is used to determine the remaining time of Green, Yellow and Red signals of traffic lights. Counting lights currently have many different types, the operating of counting down is to use 3 signals Green-Yellow-Red to calculate the time of Blue, Yellow, Red. Every time the light starts to operate, it takes 2 signal lights cycles to determine the counting time: ON TIME, calculate the Green-Yellow-Red time. Go to the next cycle and recalculate the time of Blue-Yellow-Red TIME. If the times of TIME 1 and TIME 2 of the times of Green-Yellow-Red are the same, the light starts showing the time to count down. During operation, the chip always calculates Blue-Yellow-Red time. If the signal cycle is changed (Blue-Yellow-Red time is different from the current one), the countdown lamp will stop working and continue Blue-Yellow-Red time, if there are 2 consecutive times Green-Yellow-Red time overlap, the lamp will continue to operate normally. The countdown light is not powered separately but always uses the Green-Yellow-Red signal as the source for the circuit. Because the Green-Yellow-Red signal is always forwarded and has a capacitor added to maintain the source during the signal transfer (the signal is usually 220V AC). Therefore, the current countdown light is only suitable for traffic signal point with

a fixed or little change cycle. With the signal light point has a changing cycle; it will make the countdown light not working well (sometimes turned off due to the change cycle). Hence, it is necessary to improve the countdown lights to meet the signal cycle that often changes according to the actual traffic volume on the intersection. Green-Yellow-Red time will be transmitted from the Central Processing unit to the countdown control circuit circuits via RS485 standard (each signal light point has from 4 to 16 countdown lights, the distance from the control box to countdown lights from 10 meters to 100 meters). Take steps as bellow:

+CPU sends light cycle to countdown lights, countdown lights receive information from CPU sends confirmation information (new cycle received) back to CPU.

+When the CPU receives all the confirmation information from all countdown lights, the CPU will send the ready instruction to the countdown lights.

+Countdown lights after receiving the ready command turn off the lights and wait until the signal change (from red to blue or from yellow to red) then starts displaying the countdown time of the new cycle.

+ During the operation of the system, the countdown light always sends information about its countdown cycle to the CPU, the CPU checks that information to match the calculated CPU (if it does not match, the CPU sends back repeat the countdown cycle to that light separately).

- Communication unit [5]: Using a SIM module, intergrated with GPS to get real time from satellites, transfer monitoring information, operating parameters of the system to the server, receive configuration control information which been sent from the server and text SMS to default phone number when system error.

- Display unit: Display or set values current light cycle, real time, faulty lights, control modes, warning thresholds, etc (use HMI or operate panel touch win).

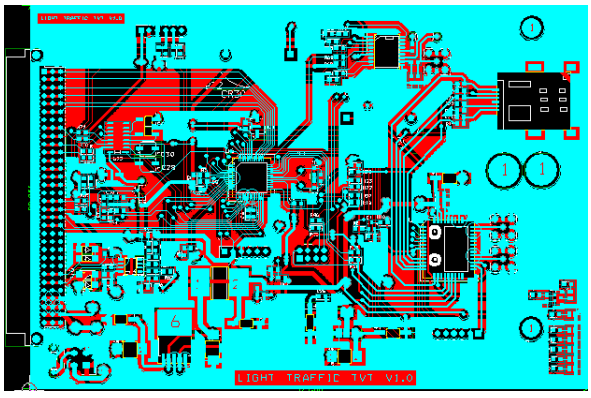


***Figure 6**. Monitoring circuit.*

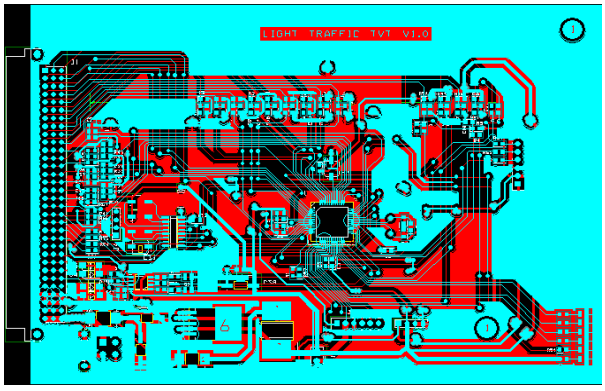


***Figure 7**. Central processing unit circuit.*

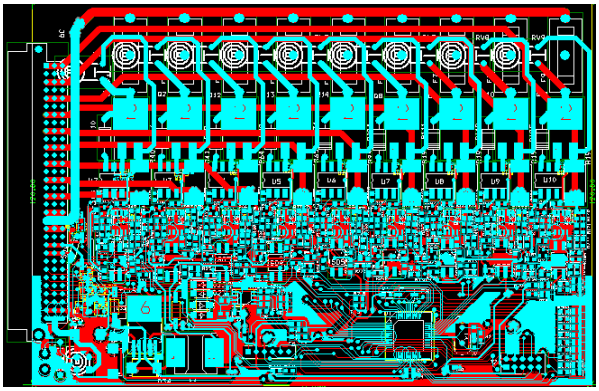


***Figure 8**. Measument circuit and lights control.*

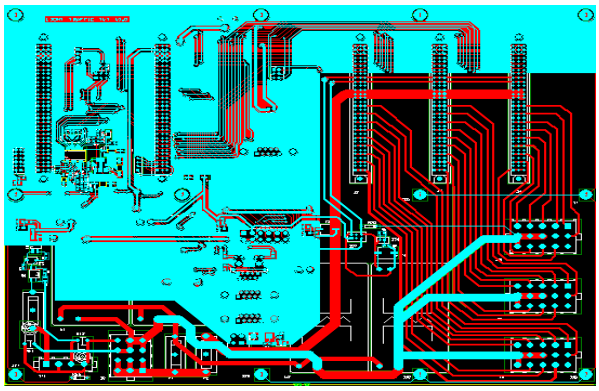


***Figure 9**. General conection circuit.*

Design the traffic data collection unit

The traffic data collection unit has the function of providing information about the density and traffic volume in the directions at the intersections to the central processing unit of the traffic light control box for calculation in order to optimal signal cycle in accordance with reality. Traffic data collection unit can use many methods such as: magnetic ring, camera, radar, infrared, etc. However, due to the traffic situation in Vietnam, it is mixed, high density and high speed, vehicle unevenness, ... so the solution of applying image processing technology obtained from the new camera results in higher accuracy than other methods.

Convolutional Neural Network (CNNs) is one of the advanced Deep Learning models that enables the construction of image processing systems with high accuracy [6]. Thus image processing algorithm application convolutional neural network performs the main function are detect group of vehicle, track splitting/merging activities of vehicle groups during movement and tracking vehicle, calculate the area of the vehicle image area, calculate the movement speed of the vehicle group. The results obtained from image processing of car groups are calculated and converted to get traffic flow data in the following direction:

- Calculate vehicle density (Df): calculate the area of the image area of each vehicle group (Ag), thereby deducing the area of occupied roads (Ao) and the converted vehicle density (DT).

- Calculate vehicle speed (Vf): calculate the speed of each group of vehicles Vg, then calculate the average speed of all groups to find the speed of the vehicle.

- Calculate the length of the queue (Lw): calculate the length of the occupied road (Lo) in the condition that the speed of the vehicle is lower than a minimum value (Vmin).

2.2.2. Use the artificial intelligence (AI) to monitor and control optimally the signal light cycle.

The roles of optimizing the monitoring and control of the signal cycle plays an important task to create an intelligent system, helping to solve existing problems on the intersection. For the monitoring part, it is necessary to evaluate the operational status of each module in the system as well as the whole system, data from measurement sensors need to be processed,

analyzed and evaluated to make early forecasts in order to promptly respond and handle to help the system always operate stably. For the control task, it needs to control accurately and optimize the signal light cycle, changing the signal light cycle flexibly to meet the actual traffic situation is important to help traffic on pint are safe and optimal, to accomplish this task, the central processing unit needs to be provided with information such as:

- Density and traffic volume on the direction of the intersections through image processing systems.

- Control modes associated with close by intersections to create the green wave (green light all the time) or setting special display modes in a short time at intersections such as: set all red light directions to disperse stuck; putting green lights in an extended direction to serve the authorities vehicles.

- Operating according to the rush hour, off peak or on holidays, working days.

- Information about the operational status of the system.

To solve the above requirements, it is necessary to apply artificial intelligence (AI) through fuzzy control theory to quantify their logical values and transfer them to the electronic brain to give signals, automatic control system. The variable of the fuzzy set is the traffic density on the directions at the intersections, from which evaluation and calculation to give the most optimal signal cycle.

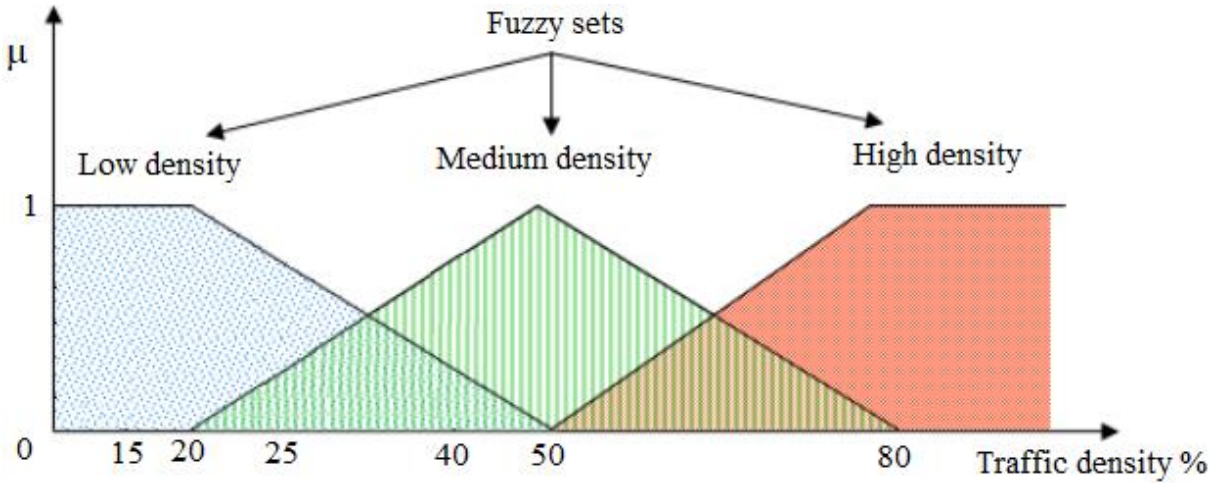


***Figure 10**. Fuzzy contrlloer digram*

In order for the system to be able to work most effectively during operation, it corresponds to each control mode such as: system status, peak, off peak hour, green wave mode, ... it is necessary to set up a separate fuzzy controller for each control mode to give the corresponding signal light cycle.

2.2.3. Control software, monitoring on server

This software system includes functional software modules such as:

- Connecting software module to control and collect data from light control boxes on the traffic signal at intersections.

- User interface software module, which allows to set the operating modes of the system, monitor the operating status of the signal lights, store and access databases, ...

- Database analysis module for evaluating light quality and group of light quality for early diagnosis of system failures, in order to promptly solve errors.

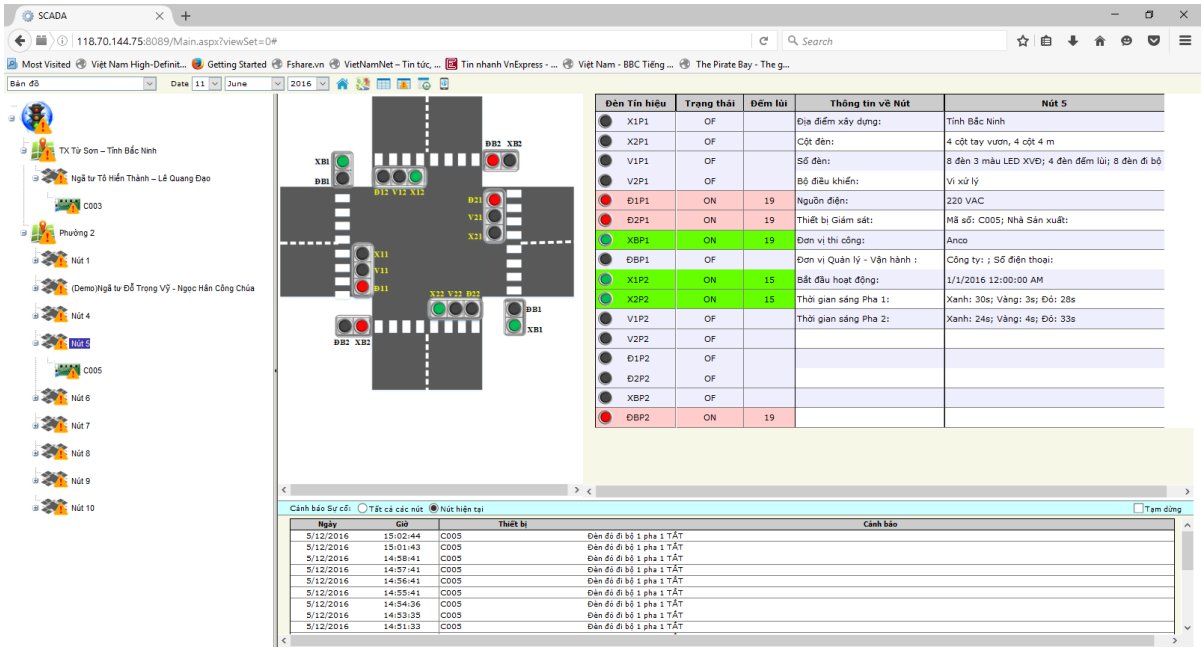


Figure 11. *Monitoring interface on server*

- Monitoring and remotely control the operation of an independent intersection or multiple intersections, pre-set the light of the day or intervene from the Operating Center.

- Connecting a continuous network of group on the route, from the Green Wave Control Center or setting special display modes in a short time at intersections such as: set all red light directions to disperse stuck; putting green lights in an extended direction to serve the authorities vehicles.

- Automatically check and ensure the safety of traffic lights, notify the Incident Center or the light signals management, easily set the display mode when there is a fault, one or several output signals are broken.

- Easily create a control phase diagram for one or more continuous intersections. Unlimited number of control lines, phase design, strategy and time cycle, scalable, adjustable and quick to set.

- The connecting camera system allows direct monitoring and supervision at intersections.

CONCLUSION

After doing the research and design of traffic light control system, it is clearly that the apply of image processing technology based on CNNs-Deep Learning combined with the use of artificial intelligence (AI) through rules of fuzzy control to design the monitoring and control system of traffic lights has brought the amazing result. It will contribute a fundermental help in reduce the accidents and traffic congestion in major cities in Vietnam.

ACKNOWLEDGMENT

In order to complete this research, the author would like to thank the management agencies and the research team of the topic "Application of image processing technology and wireless communication to design and manufacture the system intelligent monitoring and control of signal lights for a typical traffic area of Phu Ly city".

REFERENCES

[1]. Ministry of Transport, Regulation on connection of traffic light signals to road signal lights at crossroads, providing information to support the guard at intersections between roads and railways, Circular No. 28/2018/TT-BGTVT, date 14/5/2018.

[2]. Co Nhu Van and partners, Application of image processing technology and wireless communication to design and manufacture the system intelligent monitoring and control of signal lights for a typical traffic area of Phu Ly city, provincial science and technology projects, 2018.

[3]. Le Hung Lan and partners, Report on Scientific Research and Technology Development at national level, Researching and applying Information Technology and Automation in urban transport management and operation, code KC-03-21, 2005.

[4]. Le Hung Lan and partners, Report on Scientific Research and Technology Development at national level, Research, design, manufacture equipment, vehicles and systems of automatic inspection, supervision and administration for traffic safety Road, Code KC.03.05/06-10, 2008.

[5]. Co, Nhu Van; Pham, Huy Duy; Chu, Thi Thu Ha; Ngo, Thanh Binh, Design and implementation of an automatic hydrological monitoring system for hydropower plants, J. Viet. Env. 2013, Vol. 4, No. 2.

[6]. Kaiming He, Facebook AI Research Deep Residual Networks Deep Learning Gets Way Deeper.

TOWARD APPLICATION OF PIEZORESISTIVE PRESSURE SENSOR FABRICATED BY MEMS TECHNOLOGY FOR FLOOD WARNING SYSTEM IN HO CHI MINH CITY

Nguyen Tuan Khoa[1], Xuan Thang Trinh[1*], Tran Duy Hoai[1], Nguyen Thanh Phuong[1], Vu Le Thanh Long[1], Truong Huu Ly[1], Nguyen Chi Cuong[1], Hoang Ba Cuong[1], Ngo Vo Ke Thanh[1], Le Quoc Cuong[2] and Susumu Sugiyama[1]

[1]*Research Laboratories of Saigon High-Tech Park, Ho Chi Minh City, Vietnam (SHTP Labs), Lot I3, N2 Street, Saigon High Tech Park, District 9, Ho Chi Minh City, Vietnam*

[2]*Department of Information and Communications, Ho Chi Minh City, Vietnam*

[*]*Corresponding author: thang.trinhxuan@shtplabs.org*

Abstract: Piezoresistive pressure sensors were successfully fabricated by using Silicon on Insulator (SOI) 6 inches wafer. The size of dice, thickness of diaphragm, resistance of p-type piezoresistor and the range of fabricated pressure sensor were ~ 2.6 mm × 2.6 mm, ~15 μm, ~ 1300 Ω and 0-50 kPa, respectively. After packaging by using metal can package (TO-5), the pressure sensors had small offset voltage (<20 mV), high sensitivity (0.7-0.8 mV/kPa with 3V input voltage) and good linearity (< ± 1%). We have applied the pressure sensors to measure the water level on street at 3 locations (at Duong Van Can, Quoc Huong, Do Xuan Hop streets) in Ho Chi Minh City and measure tide at 1 location in Binh Duong province (in Thu Dau Mot City). Based on the obtained results, we have a plan to increase the number of monitoring locations and combine the monitoring data with Artificial Intelligence (AI) to make a flood warning system for Ho Chi Minh city.

Keywords: MEMS, Pressure sensor, Flood warning system, AI

1. INTRODUCTION

Based on several advanced properties such as low power consumption, ease of integration, low cost and fast responsibility, Micro-Electro-Mechanical Systems (MEMS) devices are considered to play an important role for the development of Internet of thing (IoT) [1]. Among MEMS devices, pressure sensor has large market and is applied in various fiels such as monitoring of fuel and measurement in automobile, air plane, blood pressure measurement and other consumer applications [2]. Pressure sensor can be also used in water level monitoring system for flood control which is one of the most serious problems in Ho Chi Minh City. In this system, both sonic sensor and pressure sensor are often utilized to measure water level to ensure the accuracy of measurement. However, the climate in Ho Chi Minh City is high humidity and rainy which cause waterlogging and disturb the measured signal of sonic sensor. Thus, in this study, we only use pressure sensor in water level monitoring system in Ho Chi Minh City.

The principle of MEMS pressure sensors are often based on piezoresistive effect in which the resistances vary with applied stress. The piezoresistive effect in Si is often larger 100-200

times to metal. In metal, the piezoresistive effect is only caused by the changing shape of the resistor. However, in semiconductor such as Si, it is mainly caused by the changing of concentration and mobility of carrier which are related to changing the lactice constant and bandgap under applied stress [3].

Piezoresistive pressure sensor often has 4 piezoresistors in Wheatstone bridge in order to cancel the effect of temperature resistivity coefficient [4]. Under applied pressure, the diaphragm will be deformed and cause stress on the diaphragm. The vary of stress leads to the vary of piezoresistance and hence, the output voltage will be change. In the piezoresistive pressure sensor, piezoresitors is often p-type Si due to its high piezoresistivity. In this sensor, the p-type piezoresistors are fabricated on n-type diaphragm to create p-n junction at the surface in order to block electronic current from top side to bottom side of the sensor which causes the noise of output signal. In p-type Si, the highest piezoresistivity is along <110> crystal direction and thus, the piezoresistors is often along <110> direction. In addition, the piezoresistive coefficient $\pi_{\ell,110}$ and $\pi_{\tau,110}$ has close value and it will help to reduce the vary of effective piezoresistive coefficient caused by misorientaion of piezoresistors during fabrication process [4].

2. EXPERIMENTAL

The piezoresistive pressure sensor (range 0-50 kPa) has 4 piezoresistors along <110> crystal direction which are arranged symmetric in order to reduce offset voltage (Figure 1).

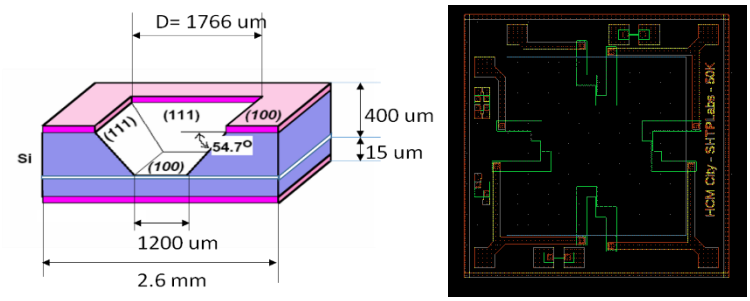


Figure 1. *Design structure and top-view of pressure sensor.*

Starting materials is Silicon on Insulator (SOI) 6 inches wafer (device layer: n-type, thickness 15 µm; SiO_2 layer: thickness 0.4 µm; and handle layer: p-type, thickness 400 µm;). The pressure sensors were fabricated by using 4 – masks processes. Firstly, four piezoresistors were fabricated on the top of device layer by Boron diffusion process using the first mask (predeposition for 30 minutes at 1000 °C, then B_2O_3 layer was removed and drive-in process was carried at 1100 °C for 30 minutes). After diffusion, resistance of the piezoresistor was measured by Jandel Multiposition Wafer Probe. The depth of diffusion was detected by staining method using $CuSO_4$ and HF [5]. Secondly, second mask was used to open contact hole. After that, 500 nm Al:Si (99% Al and 1% Si) was deposited by DC magnetron sputtering method. The third mask was used to pattern the metal. The fourth mask was utilized for etching backside using KOH solution. The diaphragm was created by KOH etching process (25% and at 75 °C) in which SiO_2 layer played as etching-stop layer in order to obtain uniform thickness of the diaphragm.

Finally, the wafer was dicing and the pressure sensors were packaged using metal can package (TO-5) (Figure 2). Fabricated pressure sensors were characterized by using Yokogawa pneumatic pressure standard (model 767402) and a home-made temperature chamber. The sensors were also measured at Ho Chi Minh City Center of Standard Metrology and Quality. For reliability testing, a home-made pressure valve was used to change pressure from 0-50 kPa repeatedly for 10000 cylces. After calibration, fabricated pressure sensors were installed in water level monitoring module.

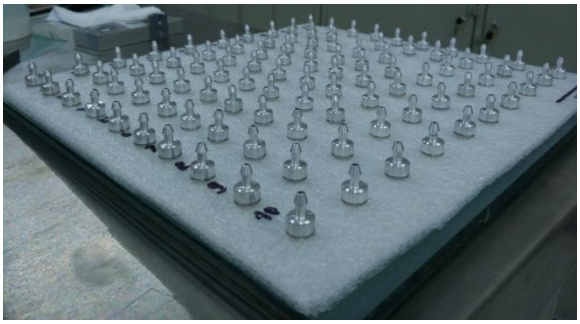


***Figure 2**. Fabricated pressure sensors.*

3. RESULTS AND DISCUSSION

After diffusion, resistance of the piezoresistor was about 1300 Ω and the depth of diffusion ~ 2.6 µm. Using 3 V input voltage, the fabricated pressure sensors had offset voltage < ± 20 mV. The offset voltage can be attributed to the mismatch of alignment during photolithography and residual stress in packaging process.

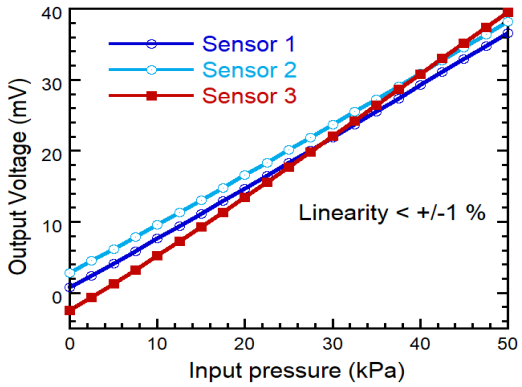


***Figure 3**. Output characterization at room temperature (25 °C) of fabricated pressure sensors.*

The span voltage and sensitivity of fabricated pressure sensors were ~ 35-40 mV, ~ 0.7-0.8 mV/kPa, respectively. From the characterization of the sensors, we obtained its linearity is smaller than ± 1%. Those parameters are similar to parameters of commercial pressure sensors. We obtained similar result from measurement of the sensors at Ho Chi Minh City Center of Standard Metrology and Quality (Figure 3).

Temperature dependence of fabricated pressure sensors was also studied in the range of 15 - 65 °C (Figure 4) which is typical for outdoor temperature in Ho Chi Minh City. The offset

voltage was increased with temperature while the sensitivity was decreased with temperature, which can be explained by temperature dependence of p-type Si resistivity.

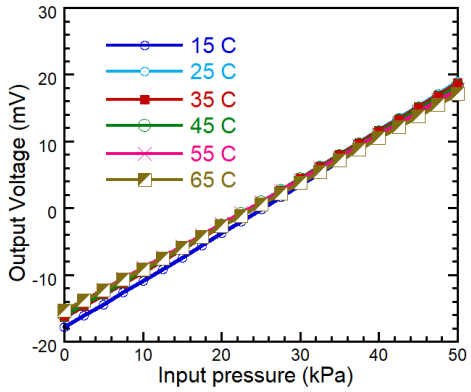


Figure 4. *Temperature dependence of the output voltage of fabricated pressure sensor.*

For reliability testing, the properties of pressure sensors remained almost unchanged after 10000 cycles of repeated pressure from 0 to 50 kPa (Figure 5). Compare to the initial span voltage, the span voltage varied < 1% during 10000 cycles.

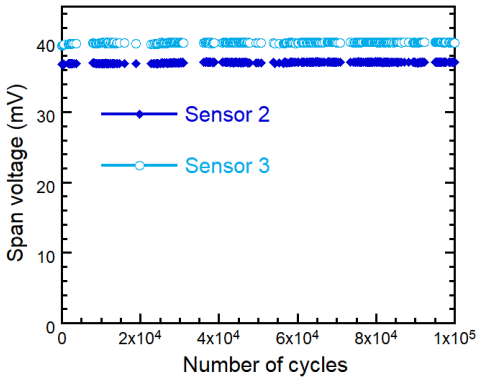


Figure 5. *Repeated load characteristics of fabricated pressure sensors.*

The fabricated pressure sensors were used to detect the changing of pressure (Δp) cause by the variation of water level on street. The relation between Δp and the height of water column (h) is expressed by following equation

$$\Delta p = h \times g \times \rho \quad (1)$$

In which g is standard gravity (~ 9.81 m/s2) and ρ is density of water (1000 kg/m3). Therefore, the height of water column (h) can be calculated by using following equation

$$h(\text{cm}) = 100 \times \Delta p\ (\text{kPa})/9.81 \quad (2)$$

We used both lithium battery and solar cell for power supply of the system to extend lifetime of the system (Figure 6). In order to reduce the noise, we used Kalman Filtering algorithm and the accuracy of water level measurement can be reduced to smaller than 2 cm.

The real-time calculated water level was sent to data center of SHTP Labs by using 3G communication technology

Table 1. *Parameters of flood monitoring device FM0919_SHTP Labs.*

Size	200 mm × 300 mm × 130 mm
Weight	350 gram
Power supply	8 V ~ 36 V
Operating temperature	0 ~ 70 ˚C
GPS	Frequency: 1575.42MHz; Sensitivity: -160dBm; Location accuracy: < 6m
GSM	Frequency: EGSM900/GSM1800MHz Sensitivity: -108dBm
Water height	Amplitude: 5 m Accuracy: ± 2 cm

We set up the water level monitoring modules using fabricated at 3 flooding locations on streets (at Duong Van Can, Quoc Huong, Do Xuan Hop streets) in Ho Chi Minh City and 1 location on river (at Thu Dau Mot City) and compared their altitudes to the standard elevation of Vietnam (VN 2000).

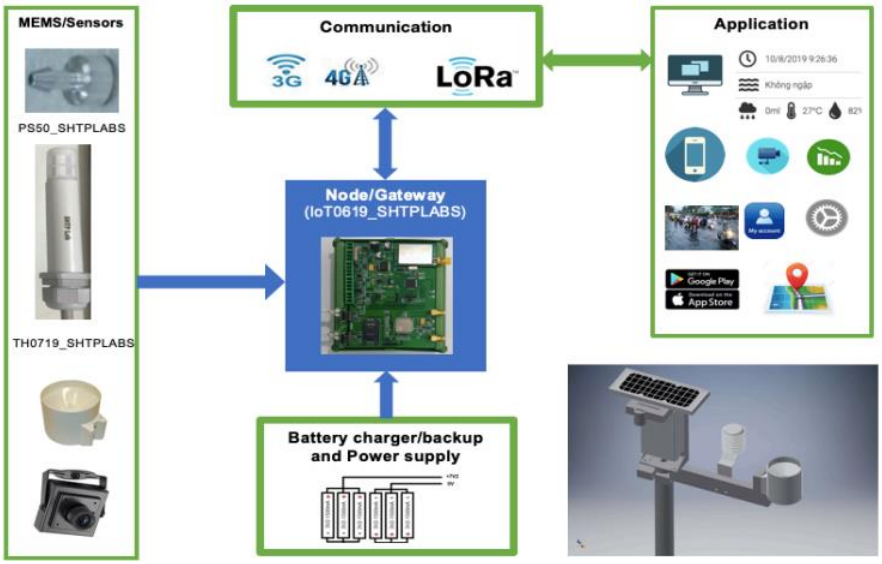


Figure 6. *Block diagram of water level monitoring system.*

In order to establish flood warning system in Ho Chi Minh City, we propose three following stages.

Stage 1: Setup water level monitoring at ~ 100 locations to monitor flooding or water level of rivers in Ho Chi Minh City. The data of all positions will be sent to our data center.

Stage 2: Using Artificial Intelligence (AI) and environmental experts to simulate flooding situations in Ho Chi Minh City.

Stage 3: Application of AI to create real-time path finding apps, SMS of flood warning for citizens.

4. CONCLUSION

Piezoresistive pressure sensors were successfully fabricated and were characterized. We applied those sensors in flood and tide monitoring system in 4 locations Ho Chi Minh City and Binh Duong. Based on the obtained results, we have a plan to scale up the number of stations to 100. Combining with AI, the monitoring data can be used to produce flood warning system for all citizens of Ho Chi Minh City.

ACKNOWLEDGMENT

This research was supported by the annual projects of The Reasearch Laboratories of Saigon High Tech Park in 2019 according to decision No. 102/QĐ-KCNC of Management Board of Saigon High Tech Park and contract No. 01/2019/HĐNVTX-KCNC-TTRD and project “Research, design, fabricate and utilize flood monitoring system at flooding places in innovative urban area in Ho Chi Minh City”

REFERENCES

[1] http://www.yole.fr

[2] Les Instituts Carnot, White Paper: Smart Networked Objects & Internet of Things, January 2010, pp. 1–8

[3] Charles S. Smith, Piezoresistance effect in Germanium and Silicon, Phys. Rev. 94 (1954), 42-49

[4] S. Santosh Kumar, B. D. Pant, Design principles and considerations for the ‘ideal’ silicon piezoresistive pressure sensor: a focused review, Microsystem Technologies, Volume 20, Issue 7 (2014), 1213–1247

[5] S. T. Ahn and W. A. Tiller, A Staining Technique for the Study of Two-Dimensional Dopant Diffusion in Silicon, J. Electrochem. Soc.: SOLID-STATE SCIENCE AND TECHNOLOGY, Vol. 135, No. 9 (1988), 2370-2373

DATA STORAGE SOLUTION FOR VIET NAM TOLL COLLECTION SYSTEM

Mai Vinh Du[1*]

[1]*Campus in Ho Chi Minh City, University of Transport and Communications, No. 450-451 Le Van Viet Street, Tang Nhon Phu A Ward, District 9, Ho Chi Minh City, Vietnam.*

[*]*Corresponding author: dumv_ph@utc.edu.vn*

Abstract: In toll collection system, clear transparency (correct and adequate), data storage solution for all transactions, vehicle images and vehicle videos to provide for post-check, audit and inspection works for operator of toll plaza and related state agencies are very important. However, at present, data storage solution for all transactions, vehicle images and vehicle videos in Vietnam toll collection system has not been uniform across the country. This paper proposed a data storage solution for all transactions, vehicle images and vehicle videos in Vietnam toll collection system.

Keywords: Toll collection system, data storage, image & video storage, NAS system.

1. INTRODUCTION

We know that, requirements on data storage for transactions, images & videos of toll collection system in circular no.49/2016/TT-BGTVT dated 30/12/2016, document no.2435/TCĐBVN-KHCN.MT&HTQT dated 27/4/2017, document no.3409/TCĐBVN-KHCN.MT&HTQT dated 9/6/2017, document no.7782/TCĐBVN-KHCN.MT&HTQT dated 7/12/2017 of the Vietnam Road Administration as following:

+ *Database storing:* Database of toll collection system must be stored and backup for all operation time of toll station (all transactions, time, type of ticket…) and pictures of vehicles from all cameras (lane camera, ALPR camera, overview camera) must be stored at least 5 years [1,4].

+ *Video storing:* The video from lane camera must be stored at least of 5 years, 480p standard (640x480 pixel), frame rate is 10 FPS. The video from cabin camera must be stored at least of 1 years, 480p standard (640x480 pixel), frame rate is 10 FPS. The video from overview camera must be stored at least of 1 years, 720p standard (1280x720 pixel), frame rate is 10 FPS [1,4].

+ *Data storage technology:* Network-attached storage (NAS) technology shall be used for database storing and video storing at all Vietnam toll stations [1,4].

In order to meet all above requirements, the paper proposed a data storage solution for all transactions, vehicle images and vehicle videos in Vietnam toll collection system.

2. CONTENTS OF DATA STORAGE SOLUTION

2.1. Network system architecture

The proposed network system architecture for data storage solution for all transactions, vehicle images and vehicle videos in Vietnam toll collection system in a toll station with 4 lanes are shown in the following Figure:

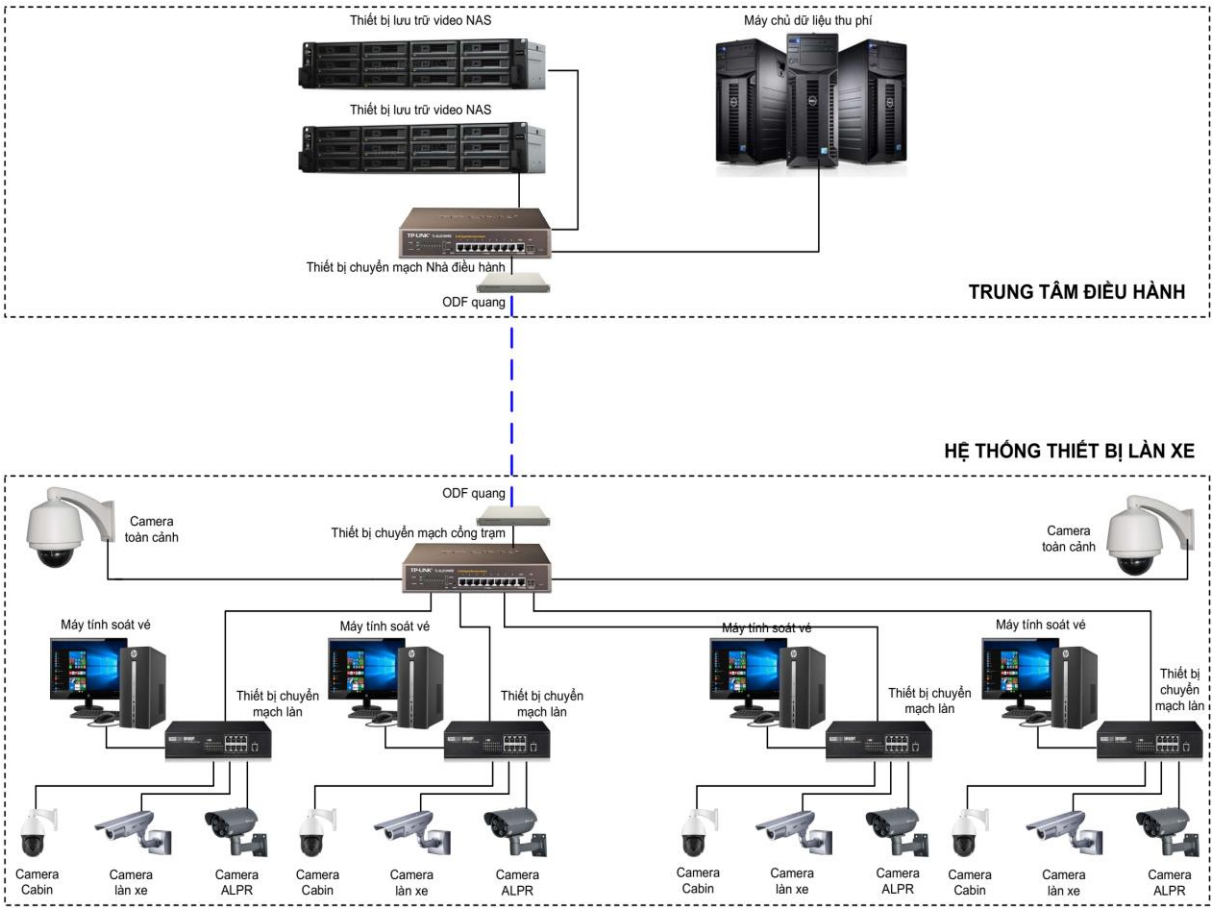


Figure 1. *Network system architecture for 4 lanes of toll station.*

NAS technology is selected in data storage solution for all transactions, vehicle images and vehicle videos in Vietnam toll collection system with many following advantages [5,6]: NAS devices are a cost-effective way to add storage space, NAS devices are easy to install and configure, configure RAID to protect stored data, NAS as a media server, avoid internet bottlenecks and outages, NAS devices are scalable, NAS devices can also be accessed remotely, NAS devices can be accessed over a wireless network, modern NAS devices are highly secure, added levels of redundancy [7,8]. Therefore, NAS technology is selected in data storage solution for all transactions, vehicle images and vehicle videos in Vietnam toll collection system. In this proposed data storage solution, we arranged the stored equipment as bellowing [5,6]:

+ Video files (cabin camera, lane camera, overview camera) shall be stored in NAS

equipment.

+ Database files (transaction data, lane camera picture, ALPR camera picture) shall be stored in Server.

2.2. Basis for data storage calculation

2.2.1 Requested parameters in circular of Vietnam Road Administration (VRA) [1,4]

+ Database files (transaction data) must be stored all operation times toll station.
+ Database files (lane camera pictures, ALPR camera picture) must be stored at least 5 years.
+ Video files (lane camera) must be stored at least 5 years, 480p standard (640x480 pixel), frame rate is 10 FPS
+ Video files (cabin camera) must be stored at least 1 years, 480p standard (640x480 pixel), frame rate is 10 FPS
+ Video files (overview camera) must be stored at least 1 years, 720p standard (1280x720 pixel), frame rate is 10 FPS
+ Standard for video compression: H.264

2.2.2 Other necessary parameters used for calculation [5,6]

+ Number of cameras (set): depend on number of cameras at toll station
+ Resolution of cameras: depend on manufacture of cameras at toll station
+ Average bandwidth/second (Kbps): depend on manufacture of cameras at toll station
+ Hard drive capacity NAS and Server: depend on design document or market
+ Volume of transactions (traffic flow)/year (KB): depend on traffic flow at toll gate
+ Data size of transactions/time (KB): depend on design document or actual data
+ Data size of transactions/year (TB): depend on design document or actual data
+ Data size of vehicle picture (KB): depend on design document or actual data
+ Data size of vehicle picture/year (TB): depend on design document or actual data

2.2.3 Formulas used for data storage calculation [5,6]

+ The formula no.1 used for video storage calculation (GB) for 1 day is as bellows:

$$\left(\left(\left(camera.no\right)\times\left(\text{second}/\text{day}\right)\times\left(Bitrate/8\right)\right)\div 1024\right)\div 1024 \qquad (1)$$

where: camera.no is number of each type of camera used in toll plaza

+ The formula no.2 used for database storage calculation (TB) for 1 year is as bellows:

$$\left(\left(\left(\left(\text{transactions per year}\right)\times\left(\text{data capacity}\right)\right)\div 1024\right)\div 1024\right)\div 1024 \qquad (2)$$

where: data capacity is capacity of one transaction data, or one lane camera picture,

or one ALPR camera picture

2.3. Data storage calculation for toll station with 6 lanes (3 lanes each direction)

2.3.1 Parameters used in calculation

Based on required parameters in Circular no.49, relevant documents, and some other necessary parameters for a toll station with 6 lanes (3 lanes for each direction), as bellows:

+ Total operation time: 16 years
+ Storage time for database files (transaction data): 16 years
+ Storage time for database files (lane camera pictures, ALPR camera picture): 5 years
+ For video files (lane camera): 5 years, 480p standard (640x480 pixel), frame rate is 10 FPS
+ For video files (cabin camera): 1 years, 480p standard (640x480 pixel), frame rate is 10 FPS
+ For video files (overview camera): 1 years, 720p standard (1280x720 pixel), frame rate is 10 FPS.
+ Standard for video compression: H.264
+ Number of cameras (set): 6 cabin cameras, 6 lane cameras, 2 overview cameras, 6 ALPR cameras
+ Average bandwidth per second (Bitrate) for 480P camera: 1024 Kbps (taken from QNAP camera)
+ Average bandwidth per second (Bitrate) for 720P camera: 1536 Kbps (taken from QNAP camera)
+ Volume of vehicle/year (total transactions): transactions in 2016 is 4,748,083 vehicles, transactions in 2017 is 5,720,699 vehicles, transactions in 2018 is 6,750,715 vehicles. Therefore, annual increase per year about 20%.
+ Hard Disk Drive (HDD) capacity of the NAS equipment: HDD 8TB
+ Hard Disk Drive (HDD) capacity of the server: HDD 2TB
+ Average capacity of transaction data/time (KB): 5KB
+ Average capacity of picture of lane camera (KB): 200KB
+ Average capacity of picture of ALPR camera (KB): 100KB

2.3.2 Video storage calculation (cabin camera, lane camera, overview camera)

❖ ***Video storage capacity of cabin camera for 1 year***

+ Apply formula no.1, we have video storage capacity of cabin camera for 1 day:

$$= \left(\left(\left(camera.no \right) \times \left({\text{second}}/{\text{day}} \right) \times \left({Bitrate}/{8} \right) \right) \div 1024 \right) \div 1024$$

= 6 *86400*1024/8/1024/1024 = 63,28 GB

+ We have video storage capacity of cabin camera for 1 year:

= 63,28 GB*365 = 23,098 GB/1024 = 22,56TB

+ Therefore, we need the number of hard drives (8TB) video storage capacity of cabin camera for 1 year are as bellow:

= 22,56/8 = 2,82 HDD (8TB)

❖ ***Video storage capacity of overview camera for 1 year***

+ Apply formula no.1, we have video storage capacity of overview camera for 1 day:

$$\left(\left((camera.no)\times\left(second/day\right)\times\left(Bitrate/8\right)\right)\div 1024\right)\div 1024$$

= 2 *86400*1536/8/1024/1024 = 31,64 GB

+ We have video storage capacity of overview camera for 1 year:

= 31,64 GB*365 = 11,549 GB/1024 = 11,28TB

+ Therefore, we need the number of hard drives (8TB) video storage capacity of overview camera for 1 year are as bellow:

= 11,28/8 = 1,41 HDD (8TB)

❖ ***Video storage capacity of lane camera for 5 year***

+ Apply formula no.1, we have video storage capacity of lane camera for 1 day:

$$\left(\left((camera.no)\times\left(second/day\right)\times\left(Bitrate/8\right)\right)\div 1024\right)\div 1024$$

= 6 *86400*1024/8/1024/1024 = 63,28 GB

+ We have video storage capacity of lane camera for 5 years:

= 63,28 GB*365*5 = 115,488 GB/1024 = 112.78TB

+ Therefore, we need the number of hard drives (8TB) video storage capacity of lane camera for 5 years are as bellow:

= 112.78/8 = 14,1 HDD (8TB)

❖ ***Total video storage capacity of 3 types of above cameras (cabin cameras, lane cameras, overview cameras):***

+ From above calculation, we have the number of HDD 8TB used for video storage capacity for 3 types of cameras (cabin cameras, lane cameras, overview cameras) are as bellow:

= 2,82+1,41+14,1 = 18,33 HDD (8TB)

+ Because the average performance of the storage device is about 80%, so the number of actually HDD (8TB) used are as bellow:

= 18,33 TB/0.8 = 22,91 HDD (8TB)

+ Therefore, the number of actually HDD (8TB) used after using about 20% for backup are as bellow:

= 22,91 TB*1,2 = 27 HDD (8TB)

+ If we choose Raid5 mode, the total number of actually HDD (8TB) used for video storage are as bellow:

= 27+1 = 28 HDD (8TB)

2.3.3 Database storage calculation (transaction data, lane camera picture, ALPR camera picture)

❖ ***Database storage capacity of transaction data for 16 years***

+ Apply formula no.2, we have database storage capacity of transaction data for 1 year (TB):

$$\left(\left(\left(\left(\text{transactions per year}\right)\times\left(\text{data capacity}\right)\right)\div 1024\right)\div 1024\right)\div 1024$$

= 6,076,714 *5*1/1024/1024/1024 = 0,028TB

+ We have database storage capacity of transaction data for 16 years (TB) with an increase of 20% each year:

= 0,028*1,2*1,2*1,2*1,2*1,2*1,2*1,2*1,2
*1,2*1,2*1,2*1,2*1,2*1,2*1,2*1,2*1,2 = 2,47 TB

❖ ***Database storage capacity of lane camera picture for 5 years***

+ Apply formula no.2, we have database storage capacity of lane camera pictures for 1 year (TB):

$$\left(\left(\left(\left(\text{transactions per year}\right)\times\left(\text{data capacity}\right)\right)\div 1024\right)\div 1024\right)\div 1024$$

= 6,076,714 *200*1/1024/1024/1024 = 1,13TB

+ We have database storage capacity of lane camera picture for 5 years (TB) with an increase of 20% each year:

= 1,13*1,2*1,2*1,2*1,2*1,2 = 8,42 TB

+ Therefore, we have database storage capacity of lane camera picture for 5 years (TB) with about 20% for backup are as bellow:

= 8,42 TB*1,2 = 10,11 TB

❖ ***Database storage capacity of ALPR camera picture for 5 years***

+ Apply formula no.2, we have database storage capacity of ALPR camera picture for 1 year (TB):

$$\left(\left(\left(\left(\text{transactions per year}\right)\times\left(\text{data capacity}\right)\right)\div 1024\right)\div 1024\right)\div 1024$$

= 6,076,714 *100*1/1024/1024/1024 = 0,57TB

+ We have database storage capacity of ALPR camera picture for 5 years (TB) with an increase of 20% each year:

= 0,57*1,2*1,2*1,2*1,2*1,2 = 4, 21 TB

+ Therefore, we have database storage capacity of ALPR camera picture for 5 years (TB) with about 20% for backup are as bellow:

= 4,21 TB*1,2 = 5,05 TB

❖ ***Total database storage capacity for 3 types of above data (transaction data, lane camera picture, ALPR camera picture):***

+ From above calculation, we have total of database storage capacity for 3 types of above data (transaction data, lane camera picture, ALPR camera picture) are as bellow:

= 2,47 TB + 10,11 TB + 5,05 TB = 17,64 TB

+ Therefore, the total number of HDD (2TB) running on Raid1 mode are as bellow:

= 17,64/2 = 9 HDD (2TB)

After summarizing the calculation results above, we have a summary table of calculation results for video storage capacity (cabin camera, lane camera, overview camera) and database storage capacity (transaction data, lane camera picture, ALPR camera picture) are shown in the following Table:

Table 1*. Summary of data storage calculation results for a toll station (6 lanes).*

No.	Content of storage	Capacity (TB)	Total capacity (TB)	Number of HDD
1	Video storage capacity of cabin camera for 1 year	22,56	146.62	27 HDD 8TB
2	Video storage capacity of overview camera for 1 year	11,28		
3	Video storage capacity of lane camera for 5 year	112.78		
4	Database files (transaction data) storage capacity for 16 years	2,47	17,64	10 HDD 2TB
5	Database files (lane camera pictures) storage capacity: 5 years	10,11		
6	Database files (ALPR camera picture) storage capacity: 5 years	5,05		

2.4. Related software for data storage solution

We know that the software of toll collection system normally has two kind of software are system software and integrated software [5,6]: The system software available in the

market such as Microsoft windows server, Microsoft windows, virus software and Microsoft office, etc. The integrated software made by contractor for specify toll collection system such as toll office system software (administration software, accounting software, supervise & post-check software, data synchronization software, database management software, other modules) and toll lane system software (ticket inspection software, overlay data software, ALPR software, other modules). In this paper, we proposed data storage solution in the toll collection system only, therefore we proposed the relating software for data storage solution only: software for data storage solution of video storage and database storage to provide for post-check, audit and inspection work of toll plaza management unit and related state agencies also, specifically as follows:

+ Video recording and overlay software: This software is compatible with cameras, and using of ticket data from SQL Server in according to the data area provided by the tollgate and integrated with existing software at the toll station.

+ Related ticketing and post-check software: provided by software at the tollgate and integrated with existing software at the toll station.

3. CONCLUSION

The paper proposed a data storage solution for all transactions, vehicle images and vehicle videos in Vietnam toll collection system, This is very important for post-check, audit and inspection works for operator of toll plazas and related state agencies in Vietnam. This solution satisfied with all requirements in related regulations in Vietnam toll collection system now. And NAS devices are a cost-effective way to add storage space, NAS devices are easy to install and configure, configure RAID to protect stored data, NAS as a media server, avoid internet bottlenecks and outages, NAS devices are scalable, NAS devices can also be accessed remotely, NAS devices can be accessed over a wireless network, modern NAS devices are highly secure, added levels of redundancy.

REFERENCES

[1] Thông tư 49/2016/TT-BGTVT ngày 30/12/2016 của Bộ Giao thông Vận tải về việc Quy định về xây dựng, tổ chức và hoạt động của trạm thu phí dịch vụ sử dụng đường bộ.

[2] Văn bản số 2435/TCĐBVN-KHCN.MT&HTQT ngày 27/4/2017 của Tổng cục đường bộ Việt Nam về việc nâng cấp hạ tầng công nghệ thông tin phục vụ sao lưu dữ liệu thu phí sử dụng dịch vụ đường bộ.

[3] Văn bản số 3409/TCĐBVN-KHCN.MT&HTQT ngày 09/6/2017 của Tổng cục đường bộ Việt Nam về việc nâng cấp hạ tầng công nghệ thông tin phục vụ sao lưu dữ liệu thu phí sử dụng dịch vụ đường bộ.

[4] Văn bản số 7782/TCĐBVN-KHCN.MT&HTQT ngày 07/12/2017 của Tổng cục đường bộ Việt Nam về việc nâng cấp hạ tầng công nghệ thông tin phục vụ sao lưu dữ liệu thu phí sử dụng dịch vụ đường bộ.

[5] Hồ sơ thiết kế thi công nâng cấp dữ liệu cho trạm thu phí cầu Rạch Miễu, Công ty TNHH tư vấn thiết kế công nghệ thông minh lập, 2019.

[6] Hồ sơ thiết kế thi công hệ thống thu phí kín (ETC, ETC+MTC) cho cao tốc Bắc Giang – Lạng

Sơn, Công ty TNHH tư vấn thiết kế công nghệ thông minh lập, 2018.

[7] J. Amol, A.R.S. Syed, Y.N. Seo, Performance Analysis of NAS and SAN Storage for Scientific Workflow, in Proceeding of 2016 International Conference on Platform Technology and Service (PlatCon), 15-17 Feb. 2016. https://doi.org/10.1109/PlatCon.2016.7456814.

[8] J. Li, Z. Wu, J.Yang, M.Xia, A secure model for network-attached storage on the grid, in Proceeding of IEEE International Conference onServices Computing, 2004. (SCC 2004). Proceedings. 2004, 15-18 Sept. 2004. https://doi.org/10.1109/SCC.2004.1358074.

A SOLUTION FOR AUTOMATIC WATERMETER READING FROM IMAGES BASED ON ROTATIONAL REGION CONVOLUTIONAL NEURAL NETWORK

Hoang Van Thong[1*], Nguyen Viet Hung[2], Nhu Van Kien[3]

[1,2] *University of Transport and Communications, No 3 Cau Giay Street, Hanoi, Vietnam*

[3] *BacBo Water Resources College, Tran Thi Phuc Road, Phu Ly, Hanam, Vietnam*

[*]*Corresponding author: thonghv@utc.edu.vn*

Abstract: In this paper, we propose an approach for Automatic Meter Reading (AMR) with water meter by leveraging the high capability of deep learning. We design a two-stage methodology, that employs the Rotational Region Convolutional Neural Networks (R2CNN) to detect region of digits and digit recognition. In the WMR literature, most datasets are not available to the research community since the images belong to a services company. As a result, in this research we made a new dataset for train and test proposal method. The results yield an impressive machine learning model and it was integrated successfully in Citywork software, which helps users verify these numbers recorded by reader and its photograph approximately.

Keywords: automatic meter reading, deep learning, convolutional neural network, recurrent neural network.

I. INTRODUCTION

Automatic meter reading refers to the automatic recording of the consumption of electric energy, gas, water, etc. for monitoring and billing. Even though there are smart meters applied for ARM, however they are not prevalent in many developing countries, where mechanical digit meters are mainly used, and meter reading is still a manual task. Some organizations require their readers to take a picture as reading proof. Since this operation is prone to errors, another operator needs to take the picture as a reading proof. This operation is prone to errors, as another operator needs to check the proof image to confirm the reading monthly. This offline checking is expensive in terms of human effort and time, and has low efficiency. Moreover, due to a large number of images to be evaluated, the inspection is usually done by sampling and errors might go unnoticed. Performing the meter inspection automatically would reduce mistakes by the human factor and save manpower.

Automatic meter reading is an application of scene text recognition. Nowadays, text recognition in the wild has received much attention from numerous researchers [1, 2], which has a variety of applications, such as automatic car license plate recognition, sign reading in autonomous vehicles, and image retrieval. There is a series of publications by a group research that focuses on this subject. Traditional approaches [3, 4] focus on the conventional Optical Character Recognition (OCR) method by first segmenting individual characters and then recognizing these characters separately. The diversity of text patterns, blurring and

backlight increase the difficulty of character segmentation. So, the performance is confined to the inaccuracy of character-level segmentation. Furthermore, recognizing each character individually ignores the relationship between the characters. Recent studies regard to scene text recognition as a sequence recognition problem without segmentation. Shi et al. [5] proposed a Convolutional Recurrent Neural Network (CRNN) to integrate CNN and RNN for text recognition. Lee et al. [6] designed an attention-based RNN approach model for OCR in the wild by weighted sequence modeling. Rayson Laroca and et al. [7] employ a smaller version of the YOLO object detector, called Fast-YOLO, for counter detection and Convolutional Recurrent Neural Network (CRNN) for digit segmentation and recognition. In this work, we propose a two-stage approach for AMR. The first stage detects a region of digits and the second stage detects digit region and recognition. We employ the rotational region convolution neuron network (R2CNN), for all phases. R2CNN is proposed by Yingying Jiang and et al. in [8]. This network is designed to detect arbitrary-oriented scene texts. We use this neural network to design an Automatic Water Meter Reading (AWMR) because retrieval of water meter images has many different directions, see Fig 1.

Figure 1. *Example of water meter images.*

One of the factors that brings success of deep learning approaches is the trained large quantity of training data in order to generalize well and yield high accuracy for unseen data. Common researches employ large public datasets to train and evaluate their models. However, these datasets were not made public in this work because they are private of companies. We have gathered 3250 images from some water supply companies in Hanoi city. We randomly split 80% total images for training phase and 20% left for testing model. Our trained model is applied as a functionality in Citywork software (see *citywork.vn*), it helps manage automatic

inspection result recording counter of operators. The Citywork is a software developed by Vietnam Citywork Software Company Limited, it helps users manage their supply water network system and customer consumption water.

The remainder of this paper is organized as follows. We briefly review R2CNN in section 2. The proposal approach is presented in section 3. In section 4 we describe experiment and result. Conclusions and future work are given in section 5.

2. REVIEW R2CNN

In [2] Yingying Jiang et al. introduce a novel framework for detecting scene texts of arbitrary orientations (see Fig. 2). It is based on Faster R-CNN [9]. The region proposal network (RPN) is used for proposing text regions and the Fast R-CNN model is modified to do text region classification, refinement and inclined box prediction. The arbitrary-oriented text detection problem is formulated as a multi-task problem. The core of the approach is predicting text scores, axis-aligned boxes and inclined minimum area boxes for each proposal generated by the RPN. To make the most of text characteristics, they did several ROIPoolings with different pooled sizes (7×7, 11×3, 3×11) for each RPN proposal and the concatenated features are then used for further detection. Their modification of Faster R-CNN by adding a smaller anchor for detecting small scene texts and using inclined non-maximum suppression to post-process the detection candidates to get the final result.

They use the RPN to generate axis-aligned bounding boxes that enclose the arbitrary-oriented texts. This is reasonable because the text in the axis-aligned box belongs to one of the following situations: *i*) the text is in the horizontal direction; *ii*) the text is in the vertical direction; *iii*) the text is in the diagonal direction of the axis-aligned box.

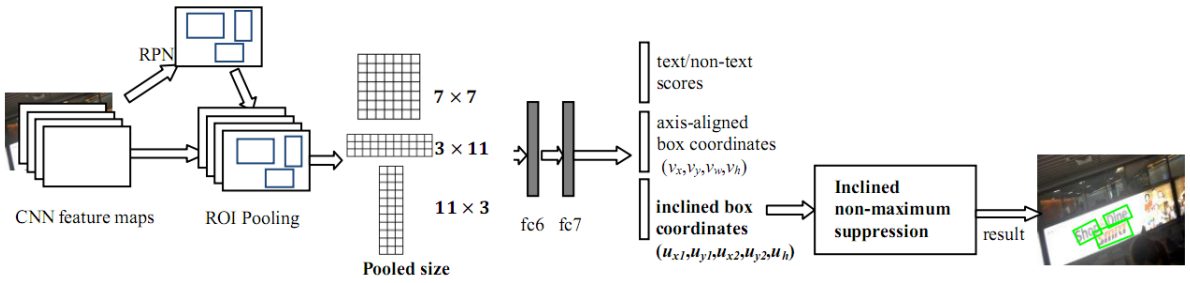


Figure 2. *The network architecture of Rotational Region CNN (R2CNN) [2].*

The Faster R-CNN framework does ROIPooling on the feature map with pooled size 7×7 for each RPN proposal. However, in fact the widths of some texts are much larger than their heights, in R2CNN so that they try to use three ROIPoolings with different sizes to catch more text characteristics. The pooled features are concatenated for further detection. Specifically, they add two pooled sizes: 11×3 and 3×11. The pooled size 3×11 is supposed to catch more horizontal features and help the detection of the horizontal text whose width is much larger than its height. The pooled size 11×3 is supposed to catch more vertical features and be useful for vertical text detection that the height is much larger than the width.

In Jiang's approach, after RPN, they classify the proposal generated by RPN as text or non-text, refine the axis-aligned bounding boxes that contain the arbitrary-oriented texts and

predict inclined bounding boxes. Each inclined box is associated with an axis-aligned box. Although their detection targets are the inclined bounding boxes, they suppose that adding additional constraints (axis-aligned bounding box) could improve the performance.

Non-Maximum Suppression (NMS) is extensively used to post-process detection candidates by current object detection methods. As they estimate both the axis-aligned bounding box and the inclined bounding box, this algorithm can either do normal NMS on axis-aligned bounding boxes, or do inclined NMS on inclined bounding boxes. In the inclined NMS, the calculation of the traditional Intersection-over-Union (IoU) is modified to be the IoU between two inclined bounding boxes.

The training loss on RPN is the same as Faster R-CNN [9]. The loss function defined on each proposal is the summation of the text/non-text classification loss and the box regression loss. The box regression loss consists of two parts: the loss of axis-aligned boxes that enclose the arbitrary-oriented texts and the loss of inclined minimum area boxes. The multi-task loss function on each proposal is defined as:

$$L(p,c,v,v^*,u,u^*) = L_{cls}(p,c) + \lambda_1 c\sum_{i\in\{x,y,w,h\}} L_{reg}(v_i,v_i^*) + \lambda_2 c\sum_{i\in\{x1,y1,x2,y2,h\}} L_{reg}(u_i,u_i^*) \quad (1)$$

λ_1 and λ_2 are the balancing parameters that control the trade-off between three terms. The box regression only conducts on text, c is the indicator of the class label. Text is labelled as 1 (c = 1), and background is labelled as 0 (c = 0). The parameter p = (p_0, p_1) is the probability over text and background classes computed by the softmax function. $L_{\text{cls}}(p, t) = -\log p_c$ is the log loss for true class c. $v = (v_x, v_y, v_w, v_h)$ is a tuple of true axis-aligned bounding box regression targets including coordinates of the center point and its width and height, and $v^* = (v_x^*, v_y^*, v_w^*, v_h^*)$ is the predicted tuple for the text label. $u = (u_{x1}, u_{y1}, u_{x2}, u_{y2}, u_h)$ is a tuple of true inclined bounding box regression targets including coordinates of first two points of the inclined box and its height, and $u^* = (u_{x1}^*, u_{y1}^*, u_{x2}^*, u_{y2}^*, u_h^*)$ is the predicted tuple for the text label. The parameterization for v and v^*given in [10], in which v and v^* specify a scale-invarian translation and log-space height/width shift relative to an object proposal. For inclined bounding boxes, the parameterization of (u_{x1}, u_{y1}), (u_{x2}, u_{y2}), (u_{x1}^*, u_{y1}^*) and $= (u_{x2}^*, u_{y2}^*)$ is the same with that of v_x, v_x. And the parameterization of u_h and u_h^* is the same with the parameterization of v_h and v_h^*.

Let (w, w^*) indicates (v_i, v_i^*) or (u_i, u_i^*), $L_{reg}(w, w^*)$ is defined as:

$$L_{reg}(w, w^*) = smooth_{L1}(w - w^*) \quad (2)$$

$$smooth_{L1}(x) = \begin{cases} 0.5x^2 & if\ |x| < 1 \\ |x| - 0.5 & otherwise \end{cases} \quad (3)$$

3. THE PROPOSED APPROACH

As shown in fig. 3, the water meter image is inputted into the first R2CNN model, the model will detect region of digits. Then the region of digits is post processed by rotating horizontally if it is not horizontal. The rotation angle is calculated based on the analysis of the digits in reverse or forward and the inclined angle of its bounding rectangle. Then the cropped region of digits is inputted into the second R2CNN to detect digit region and recognition. The single digits are arranged from left to right by x-coordinate to obtain the meter reading. In

fact, the reader only record black digits, because the red digits are not using for billing. So that, we remove red digits by colour analysis and only retrieve black digits for final results.

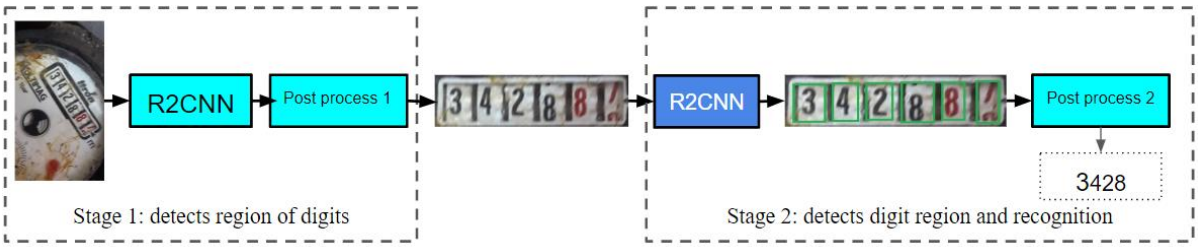


Figure 3. *Process of automatic water meter reading from images.*

4. EXPERIMENT

There is no public water meter images dataset to evaluate the performance. Therefore, we collect water meter images by mobile phones in the wild and establish a complex dataset including 3640 water meter images with six digits. These images are captured by some supply water companies in Hanoi. These images are very diverse in the horizon, vertical or inclined angle including various illumination, messy situations, complicated environment, half digit and blurring. We randomly divide 3640 images into two parts: 3000 images as training set and 640 images as test set. The training set is annotated by labelme software with 21 classes including half digits (it is a digit is changing, example 1=>2, 2=>3, etc.) and region of digits. The parameters of R2CNN is assigned as the same in [8], however, we change some parameters to be consistent with the problem data, it shows in table 1. The experiment results are presented in Table 2.

Table 1. *Some parameters of R2CNN.*

Parameter name	Value	
	Stage 1	Stage 2
IMG_SHORT_SIDE_LEN	1000	100
IMG_MAX_LENGTH	1500	400
MAX_ITERATION	500000	500000

Table 2. *The rate of accuracy on training set and test set.*

Dataset	Number of meter images	Number of meter true readings	Accuracy (%)
Training	3000	2852	95%
Test	640	593	92,6%

By analysis of the mistake results, we find that most of recognition errors appear in the case of half digits, too messy and blurring images.

Some detection results by machine learning model shown in Fig. 4 show that although the meter is inclined and messy, but the model also detects and recognizes exactly.

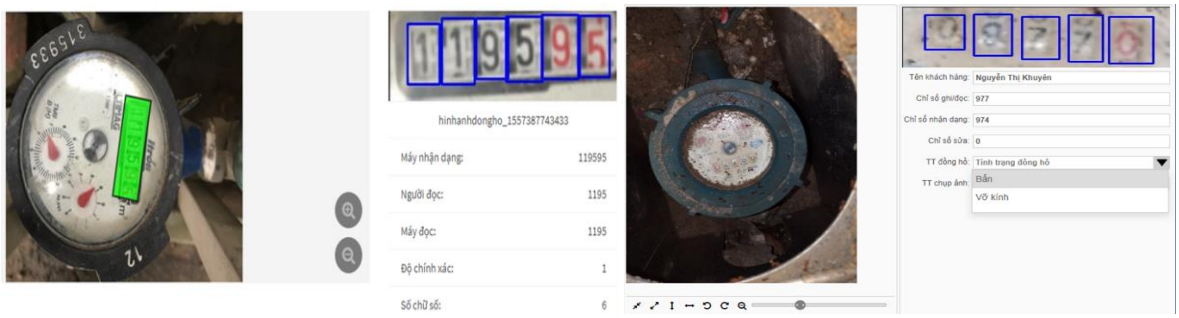


Figure 4. *Illustration of the result of digit region detection and digit recognition of the machine learning model is integrated in Citywork software.*

5. CONCLUSION

In this paper, we proposed an automatic water meter digit recognition, which is composed of digit detection and recognition. We use the R2CNN for both detection and recognition tasks, it can extract the digital region accurately and efficiently. The proposed method is trained and tested on our made dataset with 3000 images and 640 images, respectively. The experimental results show that our proposed model performs well on validity and practicability. We have applied it as an AI functionality in Citywork software.

ACKNOWLEDGMENT

This research is funded by University of Transportation and Communication under grant number T2020-CN-004.

REFERENCES

[1] Neumann, Luk´aˇs and Matas, Jiˇr´I, Real-time scene text localization and recognition, In Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on. IEEE, 3538–3545

[2] Jaderberg, Max and Simonyan, Karen and Vedaldi, Andrea and Zisserman, Andrew, Synthetic data and artificial neural networks for natural scene text recognition, arXiv preprint arXiv:1406.2227, 2014.

[3] Wang, Tao and Wu, David J and Coates, Adam and Ng, Andrew Y, End-to-end text recognition with convolutional neural networks, Pattern Recognition (ICPR), 2012 21st International Conference on, 3304–3308, 2012.

[4] Bissacco, Alessandro and Cummins, Mark and Netzer, Yuval and Neven, Hart mut: Photoocr, Reading text in uncontrolled conditions, Proceedings of the IEEE International Conference on Computer Vision, 785–792, 2013.

[5] Shi, Baoguang and Bai, Xiang and Yao, Cong, An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016.

[6] Lee, Chen-Yu and Osindero, Simon, Recursive Recurrent Nets with Attention Modeling for OCR

in the Wild. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2231–2239, 2016.

[7] Rayson Laroca,Victor Barroso, Matheus A. Diniz, Gabriel R. Gonc¸alves, William Robson Schwartz, David Menotti, Convolutional Neural Networks for Automatic Meter Reading, Journal of Electronic Imaging 28(01):1-14 (2019), DOI: 10.1117/1.JEI.28.1.013023.

[8] Yingying Jiang, Xiangyu Zhu, Xiaobing Wang, Shuli Yang, Wei Li, Hua Wang, Pei Fu and Zhenbo Luo, R2CNN: Rotational Region CNN for Orientation Robust Scene Text Detection, arXiv:1706.09579, 2017.

[9] S. Ren, K. He, R. Girshick, and J. Sun, Faster R-CNN: Towards real-time object detection with region proposal networks, NIPS 2015.

[10] R. Girshick, J. Donahue, T. Darrell, and J. Malik, Rich feature hierarchies for accurate object detection and semantic segmentation, CVPR 2014.

APPLICATION 2D HUMAN POSE ESTIMATION USING MACHINE LEARNING IN TRANSPORTATION

Nguyen Le Minh[1*], Tran Thi Dung[1]

[1]*Campus in Ho Chi Minh City, University of Transport and Communications, No. 450-451 Le Van Viet Street, Ho Chi Minh City, Vietnam.*

[*]*Corresponding author: minhnl_ph@utc.edu.vn*

Abstract: Human pose estimation from full-body images or videos is useful in many real world applications such as simulation of human activities in the game, fantasy movie, human states detection. Despite many years of research, pose estimation remains a complicated and still largely unsolved problem. Among the most significant challenges are: Variability of human visual appearance in images, in lighting conditions, variability in human physique, partial occlusions due to self-articulation and layering of objects in the scene, the complexity of human skeletal structure and high dimensionality of the pose. In this paper, we propose an effective architecture of Convolutional Neural Networks (CNNs), with an efficient object detection framework (OpenPose) and dataset COCO to build an application 2D human pose estimation in transportation.

Keywords – Deep learning, convolutional neural networks, human pose estimation.

1. INTRODUCTION

Recently, with the development of Smart Cities and IoT devices, the demand for using it in big cities is extremely large, especially in transportation. Some researches on IoT application in transportation include: Connected cars, vehicle tracking systems, public transport management or smart assistance management… At subway stations, bus stops, or at public areas, it is important to monitor human activities. More and more automatic identification devices will help gradually reduce the people served. The new idea is to build an application for the understanding of people behaviour in images and videos.

Human pose estimation with the classical approaches is using the pictorial structures framework [1]. The basic idea here is to represent an object by a collection of "parts" arranged in a deformable configuration. It should be noted, that the joints in all datasets are arranged in a tree kinematically mimicking the human body [2]. A "part" is an appearance template which is matched in an image. Springs show the spatial connections between parts. When parts are parameterized by pixel location and orientation, the resulting structure can model articulation that is very relevant in pose estimation. However, comes with the limitation of having a pose model *not* depending on image data.

In the next evolutions, most of the recent pose estimation systems have used machine learning in the development of human tracking. The idea of algorithms is estimated a 2D pose (x, y) coordinates for each joint from an RGB image. It is important to identify the joints between the limbs in the body (elbows, wrists, etc). A normal person, will be identified 15 joints on the body. Some researchers have considered these methods in the detection process.

Human Pose Estimation via Deep Neural Network - Deep Pose is the first approach developed by Alexander Toshev et al. (2014) [2]. They used model regression base on convolution neural networks (CNNs) in estimation body joints. With two datasets LSP and FLIC, focus on arm and leg. An interesting idea this model implement is refinement of the predictions using cascaded regressors, using an AlexNet backend 7 layers that output 2*k joint coordinates. Although the result is better than other state-of-art approaches, however the percentage of Correct Parts quite low, just about 0.61.

In [3], Jonathan Tompson has generated heatmaps for replace the disadvantages of direct joint regression, this approach using heatmaps for running an image through multiple resolution banks in parallel to simultaneously capture features at a variety of scales. With the purpose expected creates a discrete heatmap, a multi-resolution CNN architecture (coarse heatmap model) is used to implement a sliding window detector to produce a coarse heatmap output. However, they reuse existing convolution features, which makes this method lack structure modelling. In recent years, the next researches such as Stacked Hourglass Networks [4], Convolutional Pose Machines [5], Simple Baselines for Human Pose Estimation and Tracking [6] has continued to develop estimation through use heatmaps. Recently, researchers have started to using Deep Neural Network (DNN) to pose detection with large datasets.

In this paper, our work to propose a new effective architecture of DNNs for multi-person pose estimation. To reach this goal, we exploit the use of Openpose [8] with the COCO dataset for real-time human pose detection on the street.

The rest of the paper is organized as follows. In section 2, we describe how to build the proposed method. Our experimental results are shown in Section 3. In the final Section 4, we present our conclusions and perspectives.

2. PROPOSED METHOD

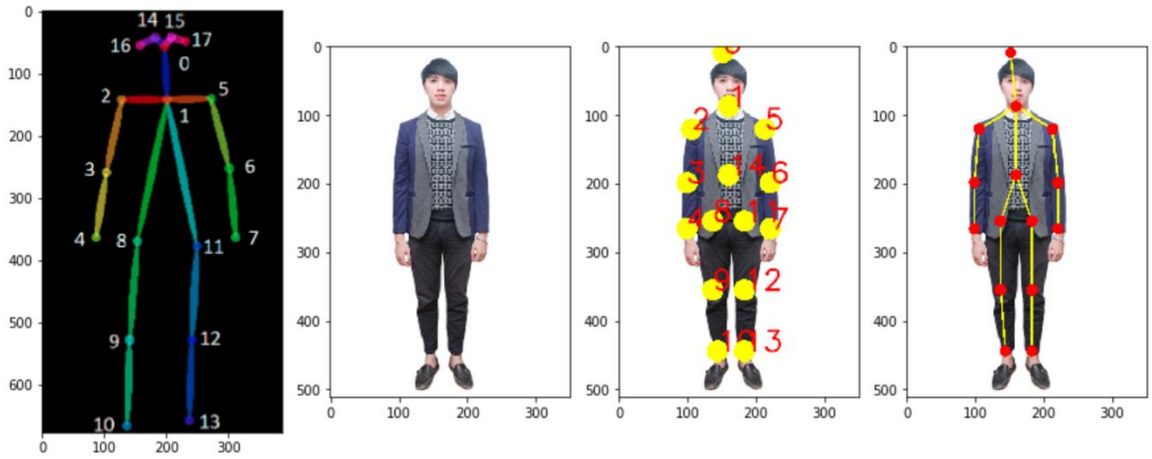


***Figure 1**. Single pose estimation using COCO key points.*

Human 2D pose estimation is the problem of localizing human body parts such as the shoulders, elbows, and ankles from an input image or video. A common approach is to follow a two-step framework which uses a human detector and solve the pose estimation for each human, in Single Person Pose Estimation, the problem is simplified by assuming the image has only one person. Multi Person Pose Estimation is more difficult because there are multiple

people in an image. The idea of this is to find the exact location of the keypoints, and then find the valid connections (or valid pairs) between the key point, assemble these connections to create skeletons for each person.

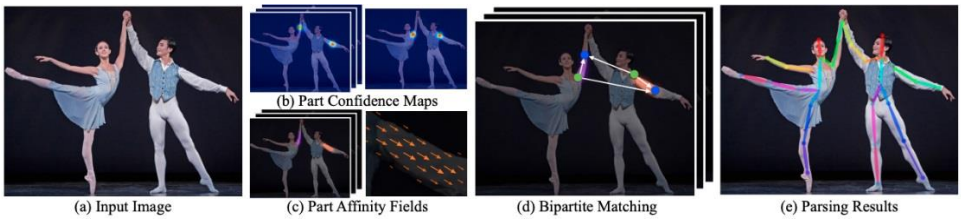


***Figure 2**. Human pose estimation method.*

Figure 2 presents the pipeline of human pose estimation method. From the figure, the detection takes place in three stages include: The first 10 layers of the VGG-19 model are used to create feature maps for the input image. Then, the feature maps are processed with multiple stages CNN to generate: (1) a set of Part Confidence Maps and (2) a set of Part Affinity Fields (PAFs) [7][9]. For (1), the image is grayscale to help identify the certain body part also known as a heatmap, there are 18 heatmaps associated with each one of the parts and indexed in the drawing of skeletons. For (2), PAF gives information about the position and orientation of pairs. With C is the number of limbs, then PAFs is the set:

$$L = (L_1, L_2, \dots, L_c) \; where \; L_c \epsilon R^{w \times h \times 2}, c \in 1 \dots C$$

Next step, a non-maximum suppression (NMS) algorithm used for detecting the parts in the image, by creating characteristic points. From this, find out the valid pair, one simple way of finding the valid pairs would be to find the minimum distance between one joint and all possible other joints. Part Affinity Maps will give the direction along with the affinity between two joint pairs, their direction should also comply with the PAF Heatmaps direction.

Finally, the confidence and affinity maps are parsed by greedy inference to produce the 2D keypoints for all people in the image. By doing the research study, we select datasets COCO to perform this idea.

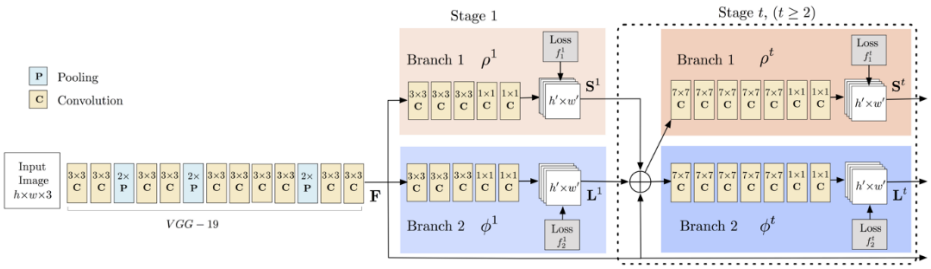


***Figure 3.** Multi-Person Pose Estimation model architecture.*

3. EXPERIMENT AND EVALUATION

A. Dataset

We perform an experiment on the well-known COCO dataset. COCO is large-scale object detection, segmentation, and captioning dataset. The COCO train, validation, and test sets, containing more than 200,000 images and 250,000 person instances labeled with key

points. COCO output format includes 18 point model trained: Nose – 0, Neck – 1, Right Shoulder – 2, Right Elbow – 3, Right Wrist – 4, Left Shoulder – 5, Left Elbow – 6, Left Wrist – 7, Right Hip – 8, Right Knee – 9, Right Ankle – 10, Left Hip – 11, Left Knee – 12, Left Ankle – 13, Right Eye – 14, Left Eye – 15, Right Ear – 16, Left Ear – 17, Background – 18.

B. Experiment setup

Our experiments have been conducted using Python programming-language with Keras, Tensorflow on Google Colab with Tesla K80 GPU.

C. Processing:

In the training process, we use the COCO dataset. After resizing the output to the same size as that of the input, then, pre-training convolutional neural network such as the first 10 layers of VGG-19, to produce a set of feature maps F.

The parsing process can be summarized into three steps:

- Step 1: Find all joints locations using the confidence maps: The goal is to create a list of joint locations of size J where each item is a list of peaks. Get the corresponding 2D heatmap for the joint in confidence maps and find the peaks by thresholding the 2D heatmap.

- Step 2: Find which joints go together to form limbs (body parts) using the part affinity fields and joints in step 1: Get all source joint peaks and destination joint peaks, create a list to store all limb connection candidates. For each source peak and each target peak: Get the direction vector, calculate the score of the current limb connection by averaging the PAFs values and add the connection to the final list.

- Step 3: Associate limbs that belong to the same person and get the final list of human poses: With the joints list from step 1 and connected limbs from step 3, find the persons that associated with either joint of the current connection

D. Result

Figure 4 shows the result of process human pose detection through the steps. From the figure, our method is able to detect multi-person from images of different postures (standing, walking, moving…) with different lighting conditions (harsh, dim).

Figure 4*. Result of process human pose detection through the steps.*

However, the result will be highly accurate when the person in the image is standing in separate positions and doesn't overlap. The difficult problem is that detection in a photo has many people, layering of objects in the scene and does not clearly show the full body. We tested over 200 videos of different human activities in the subway station, on the bus or public streets, in terms of good quality images, it has accuracy 87% with frames with less than 15 people and 78% with the opposite frames. Figure 5 shows some qualitative results of our proposed method.

Figure 5. *Result of human pose estimation before and after.*

The results also show up in different lighting conditions or with average image quality, our method still detections human pose with accuracy from 65-75% in different cases.

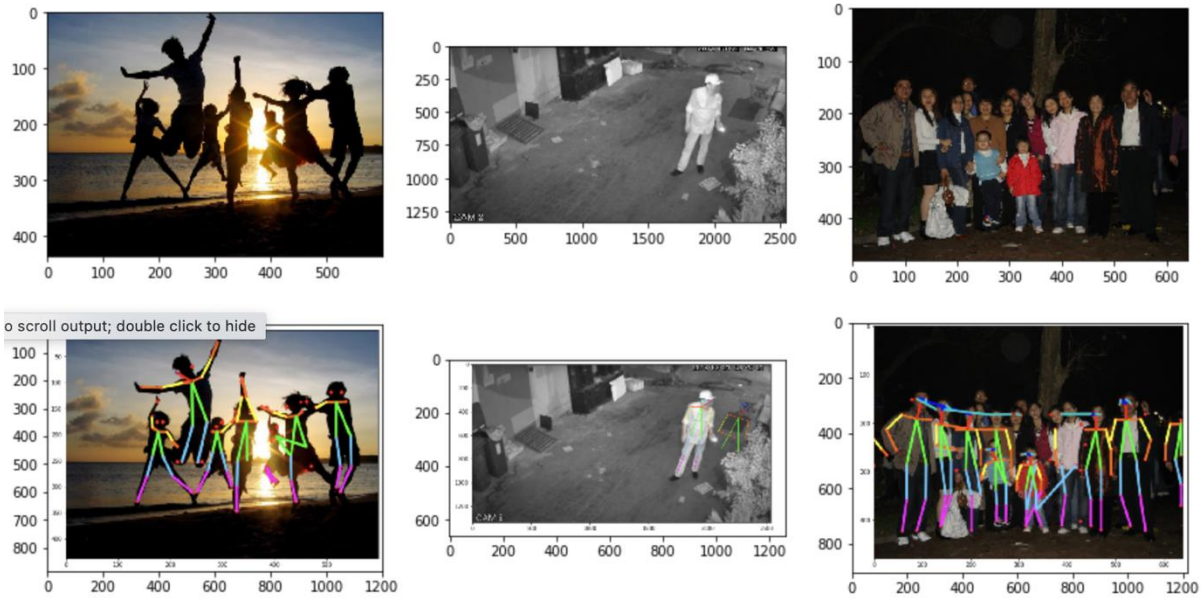


Figure 6. *Human pose estimation in different lighting conditions.*

Besides, human pose estimation also detection human activity, thereby giving emergency warnings at public areas or subway station in the future: robbery, fall, ...

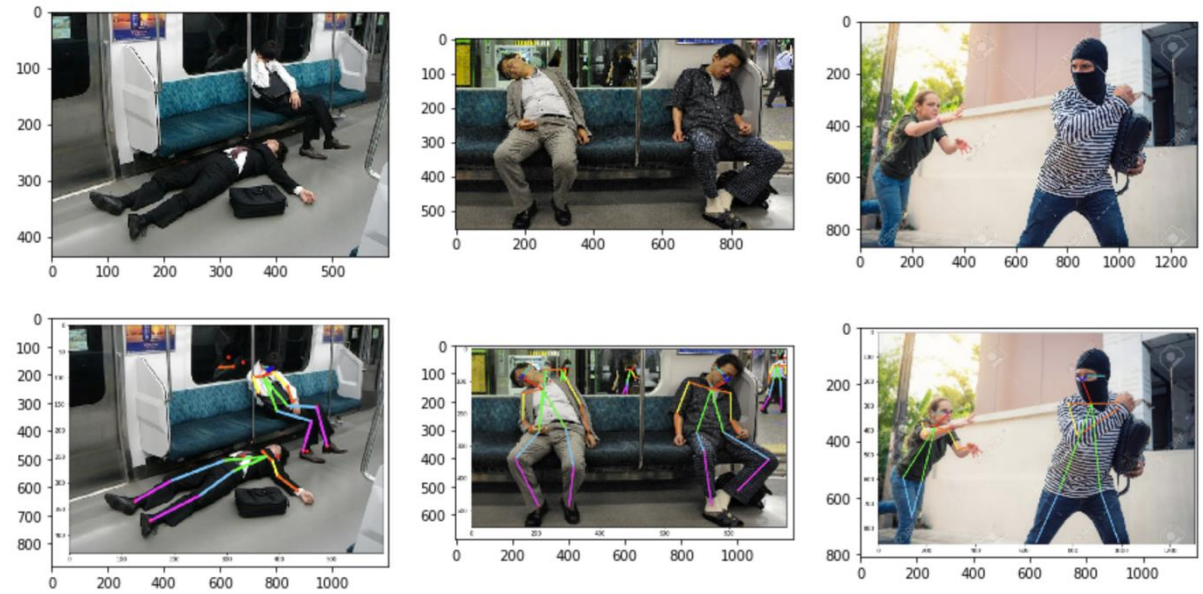


Figure 7. *Human pose estimation in human activities.*

4. CONCLUSION AND PERPECTIVES

In this paper, we built from our method using convolutional neural networks (VGG19) with the high-speed object detection framework (OpenPose) for 2D human pose estimation. The experiment on the COCO dataset shows that propose method helps in enhancing accuracy. In future work, we would like to exploit some recent effective design such as predict human posture in 3D, increasing the quality of detection processing, warning monitor human activities

in transportation.

REFERENCES

[1]. Alexander Schick, Rainer Stiefelhagen, "3D Pictorial Structures for Human Pose Estimation with Supervoxels", 2015 IEEE Winter Conference on Applications of Computer Vision

[2]. Alexander Toshev, Christian Szegedy ,"Deep Pose: Human Pose Estimation via Deep Neural Network", IEEE Conference on Computer Vision and Pattern Recognition, 2014.

[3]. Jonathan Tompson, Ross Goroshin, Arjun Jain, Yann LeCun, Christopher Bregler ,"Efficient Object Localization Using Convolutional Networks", Computer Vision and Pattern Recognition, 2014.

[4]. Alejandro Newell, Kaiyu Yang, Jia Deng ,"Stacked Hourglass Networks for Human Pose Estimation", Computer Vision and Pattern Recognition, 2016.

[5]. Shih-En Wei, Varun Ramakrishna, Takeo Kanade, Yaser Sheikh,"Convolutional Pose Machines", Computer Vision and Pattern Recognition, 2016.

[6]. Bin Xiao, Haiping Wu, Yichen Wei ,"Simple Baselines for Human Pose Estimation and Tracking", Computer Vision and Pattern Recognition, 2018.

[7]. Zhe Cao, Gines Hidalgo, Tomas Simon, Shih-En Wei, Yaser Sheikh, "OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields", Computer Vision and Pattern Recognition, 2018.

[8]. G. Hidalgo, Z. Cao, T. Simon, S.-E. Wei, H. Joo, Y. Sheikh, "OpenPose library," https://github.com/CMU-Perceptual-Computing-Lab/openpose.

[9]. A. Bulat and G. Tzimiropoulos, "Human pose estimation via convolutional part heatmap regression", ECCV, 2016.

APPLICATION OF LOGIVAN SMART TRANSPORT MODEL IN THE DEVELOPMENT OF LOGISTICS SERVICES IN VIETNAM

Van Cong Vu[1*]

[1]*Danang University of Economics, No 71 Ngu Hanh Son Street, Danang, Vietnam.*

[*]*Corresponding author: vuvc@due.edu.vn*

Abstract: Logistics is a type of service that plays an important role in promoting the country's economic growth. In addition, it is also of great significance to the transport sector in Vietnam in dealing with the challenges of traffic congestion in large cities. The paper studies the application of LOGIVAN smart transport model in developing Logistics services in Vietnam today. Research results show that LOGIVAN transport model is the optimal model in solving the problem of empty cars in the transport of goods when they go back to the place of departure. This helps minimize Logistics costs for businesses, increases income for drivers and leads towards sustainable transport development. LOGIVAN develops two platform solutions for goods owners and vehicle owners. Accordingly, the author of the article confirms the quality of this model in developing Logistics services in Vietnam and guide the operation for users via applications on personal mobile devices at the same time.

Keywords: Logistics; LOGIVAN; Vietnam.

1. INTRODUCTION

Currently, Logistics accounts for a significant proportion in the overall structure of the national economy. It plays a supporting role, connects and promotes socio-economic development of the whole country, as well as contributes to the improvement of the competitiveness of the economy. Especially, in the field of transportation, the optimization of transport and logistics development in a smart way can help reduce vehicle traffic in product distribution and solve some challenges about serious traffic congestion in some large cities in Vietnam. Therefore, Logistics is a service industry with great economic and social significance. The application of modern science and technology contributes to the increasing of labor productivity in the field of Logistics. This gradually develops and affirms the position of Vietnam Logistics service, which is a very necessary requirement.

Recently, there are many research works of domestic and foreign scientists about Logistics development and some related issues such as: *Developing Logistics service in Vietnam by Pham Trung Hai (2019); Development trend of Logistics in Vietnam in the Industrial Revolution 4.0 by Pham Hong Nhung (2019); A supply chain assessment of logistics development drivers of South Eastern Nigeria by Chinedum Onyemechi (2019); Logistics management requirements and logistics performance efficiency: the role of logistics management practices - evidence from Egypt by Hend Medhat Amin; Tamer Mohamed*

Shahwan (2020)… Works that show new research results specialize in scientific and technological solutions in developing Logistics services. However, no scientific work has clearly analyzed the status of Logistics in Vietnam from which to discuss in depth the application of LOGIVAN smart transport model in the development of Vietnam Logistics. Therefore, the author chooses a specific solution on applying LOGIVAN transport model as a research problem.

The objective of this paper is to analyze the benefits of LOGIVAN, then the author recommends applying LOGIVAN smart transport model in developing Logistics services sector in Vietnam.

The paper uses analytical research methods, aggregates secondary data source, synthesizes reviews on LOGIVAN model from experts. At the same time, the author performs operations directly on the application model, thereby, re-evaluates the problem and applies it.

2. CONTENT

2.1. Some theoretical issues about Logistics services

Recently, Logistics has gradually affirmed its important role, and has become a key economic sector in the economies of many countries. Currently in the world, it can be said that there are many definitions of Logistics.

According to the Logistics Management Board of the United States (Counseling of Logistics Management, 1991), Logistics is the process of planning, organizing, implementing and controlling effectively in terms of cost of process of material flow and storage, semi-finished products and finished products with relevant information from the starting point of the production process to the final destination for the purpose of meeting customer requirements [1].

According to the United Nations, Logistics is the activity of managing the process of transferring raw materials through the stages of storage, producing products to consumers as required by customers.

At the 7th session, Session XI, the National Assembly of the Socialist Republic of Vietnam on June 14, 2005 passed the 2005 Commercial Law, which specifies the concept of Logistics services. Article 233 - Section 4 - Chapter VI of the Commercial Law of June 14, 2005, provides for “Logistics services are commercial activities whereby traders organize one or more of the stages involved pick up goods, transport, warehousing, storage, customs clearance, other paperwork, customer advice, packaging, coding, delivery, or other related services to the goods as agreed with the customer to receive remuneration” [2].

Thus, through the definitions of Logistics, we can see that Logistics plays an important role in production, circulation and distribution of goods. It helps manufacturers to promptly and accurately capture customers' needs, boosting the flow of goods, cash flow and information from suppliers, factories, transporters, warehouses through wholesale and retail

distribution channels to end consumers more smoothly. In addition, Logistics also helps to reduce costs, satisfy customer needs and improve brand awareness of businesses.

Determining that importance, many countries already had logistics development plans and strategies. Specially, in the context of the fourth industrial revolution with more and more profound influence, Logistics in many countries has been raised to a new level. Logistics in the period of Industrial Revolution 4.0 (logistics 4.0) is a technological consolidation, eliminating the boundaries between physical, digital and biological fields, bringing a combination of virtual systems and entity. Achievements in the field of artificial intelligence (AI), integrating artificial intelligence with the Internet of Things (IoT) and modernization tools will change the whole perspective of the warehouse system and distribution of goods. In logistics activities, the Industrial Revolution 4.0 will also strengthen the connection of devices such as pallets, cranes, freight trailers ... to the Internet.

Many logistics companies around the world are starting to equip modern tools such as: Serving robots in modern warehouses; Unmanned aircraft; Autonomated guided vehicles (AGV); smart forklifts; navigation devices using WiFi, Bluetooth; Smart mobile phones participate directly in the work process; radio wave format system (RFID); barcode scanning application; inventory optimization based on cloud computing; integration of service contracts, order management, customer relations in Logistics; Blockchain technology application...

Within the scope of the article, the author mentions LOGIVAN smart transport model which is being implemented in Vietnam - the appilcation of technology to solve the problem of empty cars, contributing to reducing 30% of Logistics costs for businesses and increasing driver incomes.

2.2. Actual situation of Logistics in Vietnam today

Vietnam is considered a country with great potentials and opportunities to develop Logistics services both on land, sea and air. As a country stretching along the North-South axis, the land border is 4,550km long, bordering China in the North, Laos and Cambodia in the West [3], Vietnam has the advantage in the development of Logistics services in Southeast Asia. Specially, with the advantage of being located on the maritime transport axis, the East is bordered by the East Sea, Vietnam has favorable conditions to become a transshipment center for international marine freight transport.

In addition, Vietnam currently possesses some great potential to develop Logistics services such as: road transport infrastructure systems, airports, seaports, warehouses, trade infrastructure, logistics centers which are constantly expanded on a large and widespread scale. Besides, the accompanying services have been promptly responding to the very diverse requirements of the domestic and regional Logistics market. Customs clearance procedures for exports have also improved significantly. As a result, in recent times, along with the development of other industries in the Vietnamese economy, Logistics has also made great progress with great growth potential.

Joining the Vietnam Logistics market, according to the data published in VLA 2018 White Paper Logistics, if in 2016 the number of Logistics service enterprises in Vietnam was 22,366, by 2018, this figure would be about 30,971 enterprises respectively, an increase of 30% [4]. Currently, 30 transnational logistics service providers are operating in Vietnam with big names such as: DHL, FedEx, Maersk Logistics, APL Logistics, CJ Logistics, KMTC Logistics… [5]. By business sectors, Logistics enterprises in Vietnam are classified into 4 types.

Table 1. *Classification of Logistics service enterprises by business sector.*

Number	Type of business
1	Transport operation enterprises: Transportation services (road, sea, air)
2	Enterprises exploiting infrastructure at the nodes *(ports, airports, stations ...)*
3	Enterprises exploiting loading and unloading warehouses and Logistics services
4	Freight forwarding service business, transport agents, shipping agents, customs clearance agents, 3PLs and other businesses such as Logistics software solutions, consulting and inspection, financial checks

Source: *Author synthesized according to Vietnam Logistics Report 2017 – Ministry of Industry and Trade*

In recent years, Vietnam's logistics capacity has gradually been improved and achieved many positive results. According to statistics of the World Bank (World Bank) in 2018, Vietnam's logistics industry currently has a growth rate of 16% per year and is expected to continue to grow in the coming years. LPI report *(Logistics Performance Index - the index evaluating Logistics performance)* in the publication "Connecting to complete 2018" of World Bank in 2018, Vietnam's ranking increased by 25 places to 39/160. Accordingly, Vietnam surpassed at 39th rank with a significantly improved LPI score of 3.27, the highest in 6 rankings, ranked 3rd place in ASEAN, after Singapore ranked 7th and Thailand ranked 32th. Average LPI of Vietnam over the years from 2012, 2014, 2016, 2018 World Bank reports is currently ranked 45/167 countries [6]. According to the Vietnam Association of Logistics Services Enterprises (VLA), the scale of Vietnam's logistics market has been constantly increasing with the rapid growth of import and export turnover over the past years, reaching about 40-42 billion USD/year. In 2019, Vietnam's logistics service has a relatively high growth rate, reaching 12-14%; the percentage of enterprises outsourcing Logistics services is about 60-70%, contributing about 4-5% of GDP. The Government set a target that by 2025, the proportion of logistics service industry in GDP will reach 8% -10%, the growth rate of services will reach 15% -20%, the proportion of outsourcing logistics services will reach 50%. -60%, Logistics costs go down [7].

Determining the important role of Logistics in the national economy in general and in the field of Transport in Vietnam in particular, the Government has issued a number of legal

documents clearly showing its commitment to support and create a favorable environment for the purpose of improving competitiveness and developing Vietnam Logistics service. On January 22, 2014, the Prime Minister issued Decision No.169/QD-TTg approving the Project on Development of Logistics Services in Vietnam's Transport Sector to 2020, with orientations to 2030. Accordingly, the general purpose of the project is developing Logistics services in Transport field, in order to enhance economical effect in Transport operation; contributing to the success of Vietnam Sea Strategy by 2020. On February 14, 2017, the Prime Minister issued Decision No.200/QD-TTg on Approving the action plan to enhance Vietnam's competitiveness and develop Logistics services to 2025. Among them, 6 viewpoints on developing Logistics services of the Decision are: (1) Logistics is an important service industry in the overall structure of the national economy; (2) Logistics becomes a service industry that brings high added value; (3) Developing a healthy Logistics service market; (4) Maximize the advantages of strategic geographic location to become an important Logistics hub in the region; (5) Improve the competitiveness of enterprises providing Logistics services. To develop enterprises providing Logistics services in terms of quantity, size, human resource, and high competitiveness in domestic and international markets; (6) The State plays a supporting role, creating a favorable environment. On July 6, 2018, the Prime Minister signed Decision No.27/2018/QD-TTg promulgating the Vietnam Economic Sector System and this is the first time Vietnam has its own Logistics industry code *(Code 52292: Logistics)*. On July 18, 2018, the Prime Minister signed the Directive No.21/CT-TTg on promoting the implementation of solutions to reduce logistics costs, effectively connect transport infrastructure...

Thus, the close attention and appreciation of the Government to the Logistics service industry shows that Logistics is a very important sector of Vietnam's economy. And both short-term and long-term strategies are required to develop Logistics services in the coming time. Inheriting the development of science, technology, application of artificial intelligence in the field of Logistics, transport optimization is also an important factor contributing to the development of Logistics services in Vietnam.

2.3. Applying LOGIVAN smart transport model in developing Logistics services in Vietnam today

2.3.1. LOGIVAN application

LOGIVAN application is a technology solution that provides transport services 4.0 based on the direct connection of goods owner and empty truck network nationwide. LOGIVAN application helps businesses search and manage trucks online. At the same time, it helps truckers increase their income by up to 30%.

Established in September 2017, LOGIVAN's mission is to digitize the trucking industry, providing a high-quality automated integration system that saves logistics costs for all Vietnamese businesses. In March 2018, LOGIVAN won the Uber Exchange start-up contest, the only start-up to win a ticket to Silicon Valley, USA and received direct advice from UBER Global Chief Technology Officer - Thuan Pham. By April 2018, LOGIVAN

successfully raised US $600,000 from Insignia Ventures Partners. It began to expand operations in Ho Chi Minh City and other major economic centers in Vietnam, serving companies with high transportation needs. In July 2018, LOGIVAN won the Best Startup Company in PITCH 2018 at RISE 2018. Overcoming more than 800 teams from around the world, LOGIVAN won the first victory for Vietnamese startups at the Asia Technology Summit. In August 2018, LOGIVAN invested an additional $1.75 million in Series A round led by Ethos Partners, Insignia Venture Partners and VinaCapital Ventures. From here, LOGIVAN continued to upgrade technology products to serve the needs of freight business and facilitate transparency in truck tracking and management. In December 2018, LOGIVAN won the First Pitch@Palace Global 2018 organized by Prince Andrew - Duke of York. In addition to providing a breakthrough solution in the field of Logistics, LOGIVAN is also praised by the Award Council for having a positive impact on the society, economy and living environment in Vietnam. In February 2019, LOGIVAN successfully called for US $5.5 million from Alpha JWC and other reputable Asian investors. With this investment, LOGIVAN aims to analyze and integrate data to increase supply chain efficiency, expanding the network of 40,000 transport partners with hundreds of orders per day.

In mid-2019, the Prime Minister's Decision No.703/QD-TTg on the application of transportation technology to improve service quality and reduce logistics costs in total GDP has become a driving force for LOGIVAN to continue proving its role. LOGIVAN has launched an *apple (automatic pricing function for LOGIVAN: Code E), reel (a system suggested by the field of machine learning based on artificial intelligence)* pricing system and integrated into the 5-star LOGINOW package. On October 21, 2019, LOGIVAN officially launched the LOGINOW booking service, which helped shippers get instant quotes and 100% have cars and many other outstanding utilities. With LOGINOW, the goods owner will be directly quoted and have drivers arranged for the order by LOGIVAN. LOGINOW rates are not only displayed immediately for all routes, but also good and competitive in the market thanks to LOGIVAN's modern quotation system. From there, it helps shippers look up prices quickly and accurately with the best freight for all shipping routes, contributing to saving time and reducing costs for businesses' Logistics.

Thus, with the application model of LOGIVAN technology, the goods owner will always find the most suitable car, no more complicated and time-consuming steps to find cars like the traditional way of finding cars before. There is no longer a situation where a car is canceled or a vehicle cannot be found. At the same time, LOGIVAN's staff will monitor and manage the entire shipment's itinerary. At the same time, the goods owner can also know the status of their goods by actively tracking the journey of vehicles displayed online on the LOGIVAN Goods owner application.

Specially, LOGINOW also has cargo insurance and contracting services to bring transparency and peace of mind to goods owner. Therefore, goods owner will no longer worry about the safety of goods when using the LOGINOW feature.

All shipments booked through the LOGIVAN application and LOGINOW usage are managed by a 5-star quality driver team selected by LOGIVAN. Only drivers with full

information, professional working attitude, extensive shipping experience can receive LOGINOW orders. This helps minimize the risks when transporting goods, bringing peace of mind for shippers.

2.3.2. Evaluate the effectiveness of the LOGIVAN model

The author's research results show that, LOGIVAN transport model is a smart and efficient transport model, users will gain the great following benefits:

Firstly, minimizing logistics costs: LOGIVAN takes care of the whole process of searching for freight vehicles and transportation process. Businesses will save labor costs and simplify the transportation process.

Secondly, the benefits from LOGIVAN come from eliminating intermediaries in the supply chain, connecting goods owner' needs directly to sources of empty vehicles nationwide and helping vehicle owners increase their monthly income. The transport connection application is completely free. Drivers do not have to pay any costs when using LOGIVAN - Vehicle owners. Supporting quick orders, it is rated for increasing 30% income for drivers. If each delivery truck is ordered in the traditional way, you are paying 30% of the money for a service you do not use. It is a fact that not many people know. 30% of this cost is actually paid for the way back to the place of departure of the car because at that time the car was empty without goods. It is a serious waste for a car to return to its starting point without carrying any cargo. The car still consumes a significant amount of gasoline but does not optimize the transport capacity. Using LOGIVAN - transportation model applying technology and artificial intelligence in operation has become a "dual" solution. The goods owner do not have to pay 30% of the cost of turning and drivers take advantage of empty cars to return to increase income.

Thirdly, goods owner easily rent a cargo truck with just a few simple steps. Goods owner are quoted directly on the app and the costs are less than the market price. The vehicle owner can proactively negotiate shipping costs with the goods owner.

Fourthly, ensuring the safety: the LOGIVAN application displays the entire route map of the cart. Businesses easily grasp the status of their orders, limit late deliveries and worry about goods.

Fifthly, ensuring transparency: All transport activities are carried out by contract and full documentation. In addition, LOGIVAN still has the form of monthly debit to eliminate cumbersome payment process.

Sixthly, with the breakthrough idea to take advantage of the empty direction of the vehicle, LOGIVAN contributes to reducing emissions from vehicles. As a result, it reduces the harmful effects of greenhouse gases and environmental pollution throughout the country.

In addition, using the service of LOGIVAN also helps to limit the amount of traffic on the road, avoiding traffic jams and accident risks. If LOGIVAN application is widely implemented, it will promote Vietnam's shipping industry in a sustainable manner.

Evaluating LOGIVAN application, some LOGIVAN logistics experts and big customers have many positive comments. Ms. Bui Truc Quynh, Director of HP Agriculture Company responded after using LOGIVAN: *"LOGIVAN makes our transportation much more and more flexible, while ensuring quality and progress"*. Ms. Do Phuong Thao, Director of Thao Nguyen Company - Top 3 cashew exporters, top 500 large enterprises in Vietnam, said: *"LOGIVAN is very easy to use and can find a truck much faster than calling the transportation parties"*. Mr. Ta Quoc Loi, Director of Logistics of Miwon Vietnam Co., Ltd. evaluated *"LOGIVAN works professionally and youthfully. Miwon always receives support and coordination of LOGIVAN when arising, even on holidays"* [8].

Thus, it can be said that LOGIVAN is an effective and suitable application for users in Vietnam.

2.3.3. Manipulate LOGIVAN application

LOGIVAN develops two transport technology products, LOGIVAN - Goods owner and LOGIVAN - Vehicle owner. Corresponding to the two technology products, the user is the owner of the goods and the owner of the truck will use two applications that suit his needs through mobile devices. In CH Play app of android devices, users can easily download the application. The interface of the application is simply identified.

Figure 1. *LOGIVAN application interface – Goods owner and LOGIVAN - Vehicle owner.*
Source: *Author made.*

2.3.3.1. For goods owners

After downloading the *LOGIVAN Chủ hàng* application to the machine, the owner of the goods proceeds to create a personal account. The steps to create an account are performed sequentially with simple operations: The owner only needs to register the current phone number, then enter full name then the account will be established.

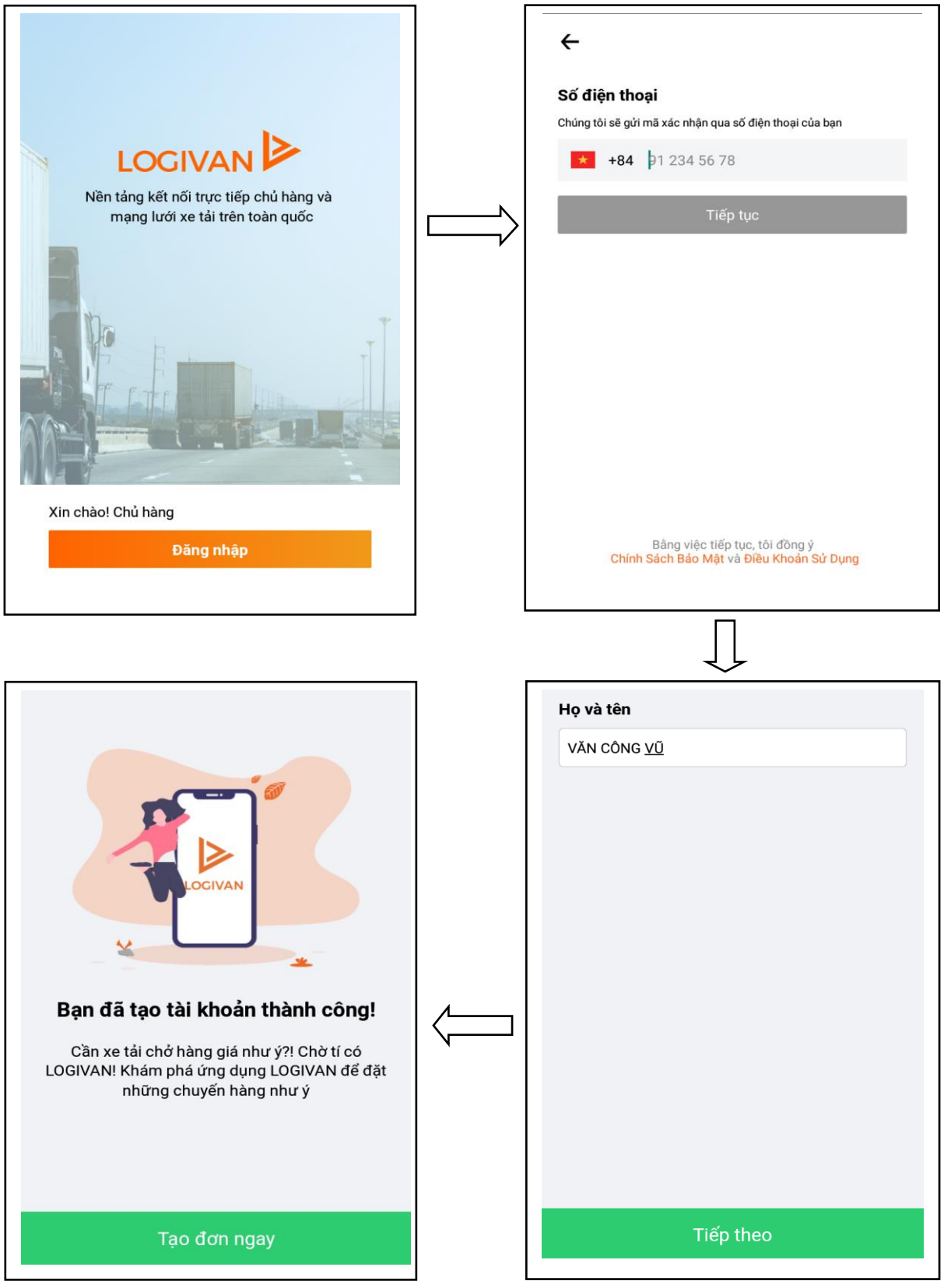


Figure 2. *Register account for commodity owner. (**Source:** Author made).*

After creating an account, the owner of the commodity can begin working on his or her transport needs. By entering the information: Shipping address (including the address of the goods to be transported and the address of the destination); Loading time; Name of goods and quantity of goods. The application will immediately display the best shipping price for owners. The price that the owner of the goods must pay is always the best price because this is a transport service on the return car, instead of no goods on the way back.

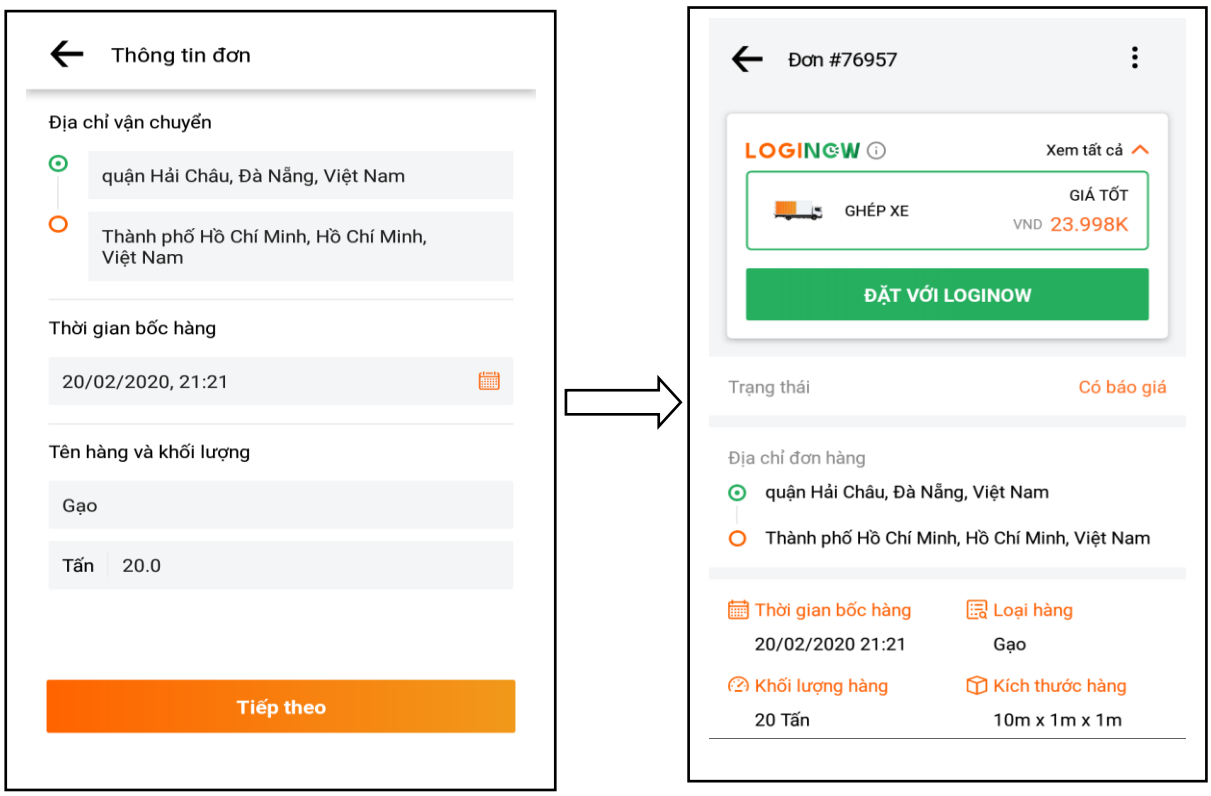


Figure 3. *The owner of the goods carries out the service on demand.*
Source: *Author made.*

After the completion of the order, if the owner of the goods wants to cancel the order for an objective or subjective reason, there is no need to continue using the service. It is possible to cancel an order by choosing one of the following reasons: No longer in need; order information changed; another car found; the type of vehicle or driver is inconsistent with the previous confirmation; the driver requires additional fees not included in the agreement; driver's attitude; driver arriving late at loading point; failure in communicating with the driver or distrust of the driver.

2.3.3.2. For drivers with vehicles involved in transportation services

The driver will download the *LOGIVAN Chủ xe* application to the phone, then proceed to register the user account. The account registration for the vehicle owner is more complete and more informative than for the owner of the goods with the purpose of creating trust with customers and management from LOGIVAN company. The driver will provide the following information: Phone number; first and last name; parking address; provinces that run often;

Identity card; business license; transport license.

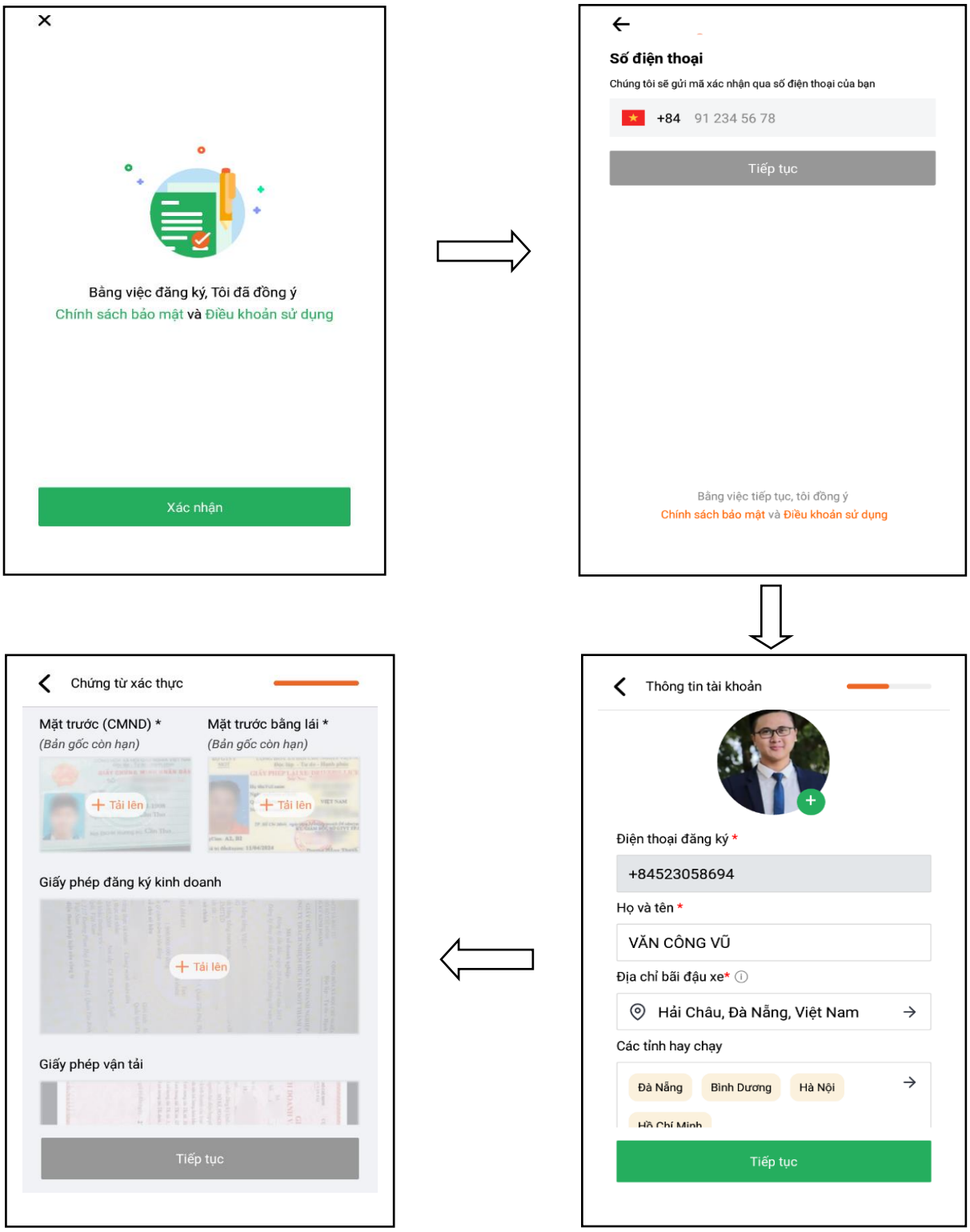


Figure 4. *Account registration operation for Vehicle Owner.*
Source: *Author made.*

After signing up for an account, the driver can start accepting shipments according to customer needs. Orders will appear on the vehicle owner's app with information about the location, time, type of goods, customer contact information. With the policy and model of LOGIVAN, the vehicle owner has the right to receive or not to receive shipments to suit his transport needs. If the driver accepts the order, he or she will announce the fee to customers.

Figure 5. *The vehicle owner performs the return freight service.*
Source: *Author made.*

After the quotation, if the customer agrees, the truck owner will conduct some follow-up operations to receive the order, when receiving and loading the vehicle, the vehicle owner click on the item: *"Tôi đã bốc hàng"*. At the end of the transaction, after the vehicle owner has delivered the goods to the customer, the vehicle owner clicks the item: *"Tôi đã dỡ hàng"* and completes the transaction. The implementation of these operations is to help LOGIVAN's staff keep track of the status of orders to update customers, creating the reputation and professionalism of the service.

Figure 6. *Vehicle owner receives order status. (**Source:** Logivan).*

3. CONCLUSION

LOGIVAN intelligent transport model is an innovative transport model of Vietnamese people, aiming at the goal of transport optimization, bringing sustainable development to the national transport network.

With the development of services on two fundamental solutions: LOGIVAN Chủ hàng and LOGIVAN Chủ xe, LOGIVAN has solved the problem of empty return for businesses, saving costs for customers and increasing income for vehicle owners. Simple to use and performed on personal mobile devices, LOGIVAN ensures maximum convenience for users. Applying LOGIVAN smart transport model is a new direction, consistent with the development of Logistics services in Vietnam, bringing high economic efficiency, contributing to encourage businesses to innovate and apply science and technology in production and business.

However, the development of Industry 4.0 requires the LOGIVAN team to constantly research and innovate services to suit social needs in certain stages. In addition, international cooperation to improve technology, scale and market operations are important requirements in the coming time.

REFERENCES

[1]. Donal F. Wood, Logistics Business, https://www.britannica.com/topic/logistics-business#ref528537

[2]. National Assembly, Commercial Law, Hanoi, 2005. https://moj.gov.vn/vbpq/Lists/Vn%20bn%20php%20lut/View_Detail.aspx?ItemID=18140

[3]. Government Portal of the Socialist Republic of Vietnam, http://chinhphu.vn/portal/page/portal/chinhphu/NuocCHXHCNVietNam/ThongTinTongHop/dialy

[4]. Ministry of Industry and Trade, Vietnam Logistics Report 2018, Hanoi, 2018.

[5]. T.H. Pham, Developing Logistics services in Vietnam, Finance Magazine, April 28, 2019. http://tapchitaichinh.vn/nghien-cuu-trao-doi/phat-trien-nganh-dich-vu-logistics-tai-viet-nam-306129.html

[6]. World Bank, 2018, https://openknowledge.worldbank.org/bitstream/handle/10986/29971/ LPI2018.pdf

[7]. Chung Thanh, 70% of logistics enterprises are concentrated in Ho Chi Minh City and surrounding areas, Ho Chi Minh City Government Portal, August 1, 2019. http://tphcm.chinhphu.vn/70-doanh-nghiep-logistics-tap-trung-o-tphcm-va-vung-lan-can

[8]. Website: https://www.logivan.com/

550 – BINH DUONG OVERPASS USING BUILDING INFORMATION MODELING

Huynh Xuan Tin[1,2*], Nguyen Quoc Chuong[2], Vo Phu Toan[2]
Do Minh Truyen[2], Nguyen Truong Giang[2], Nguyen Xuan Viet[2]

[1]*Campus in Ho Chi Minh City, University of Transport and Communications, No. 450-451 Le Van Viet Street, Tang Nhon Phu A Ward, District 9, Ho Chi Minh City, Vietnam.*

[2]*V7 Engineering and Technology Company Limited CO.LTD, No 448/5E Le Van Viet Street, Tang Nhon Phu A Ward, District 9, Ho Chi Minh City, Vietnam.*

[*]*Corresponding author: tinhx_ph@utc.edu.vn*

Abstract: Artificial intelligence (AI) has become the dominant technology to improve productivity in the construction sector thanks to the development of Big Data in construction. The application of AI and Big Data in the transport sector is expected to help Vietnam's transportation industry to address challenges and towards sustainable transport development. Building information modelling (BIM) is proved effectiveness in building and management information during project life cycle. In this article, the information model of 550 – Binh Duong overpass composite by box steel girder and reinforced concrete slab will be introduced.

Keywords: AI (Artificial Intelligence), BIM (Building Information Modelling), BIM data, construction engineering, Level of Development (LOD).

1. INTRODUCTION

Building information modelling (BIM) has become mainstream technology construction in recent years. BIM can apply many fields such as design, construction and project management to solve the problems of material wastage, low productivity, and poor information exchange. BIM defines all property information related to a full lifecycle of a target structure, from planning and design to construction, operation, and maintenance, in terms of technology of construction, production, management, and utilization. A BIM model, generating by such technology, is a virtual structure built with 3D graphic data and non-graphic data or property information. BIM helps to decrease from 5% to 20% initial investment cost and about 30% operation and maintenance cost [1]. This article presents BIM in design for 550 – Binh Duong overpass composite by box steel girder and reinforced concrete slab. The bridge is simulated of the full range 3D information models (LOD 350) for virtual reality and material quantity management. The project BIM model helps to control the detailed cost, safely construction and effective operation.

2. 550 – BINH DUONG OVERPASS COMPOSITE BY STEEL BOX GIRDER AND REINFORCED CONCRETE SLAB USING BIM

2.1. Design solutions using 3D models

BIM model opens up new perspectives on a project life cycle management. Revit software, an Autodesk product, used for the model. 3D information model of 550 - Binh Duong overpass (Figure 1) includes 5 spans of 40m length and 16m width, steel box girders, and concrete slab. The model, basing on 2D drawings, is used for quantities extraction, construction bidding and then for construction and operation.

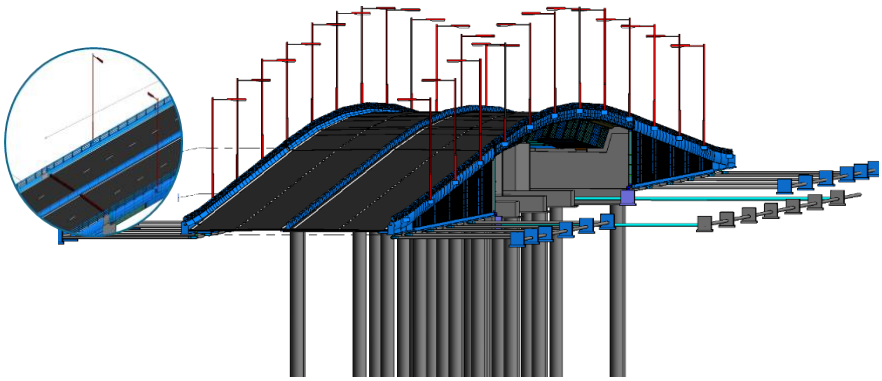


Figure 1. *3D model 550 – Binh Duong overpass using Autodesk Revit.*

The 3D model also helps to check and modify errors of the design process. In Figure 2, design errors were found in the detailed preliminary design for the pier's structure, which can be directly modified in the 3D model with the exact coordinates, dimensions, and properties [2].

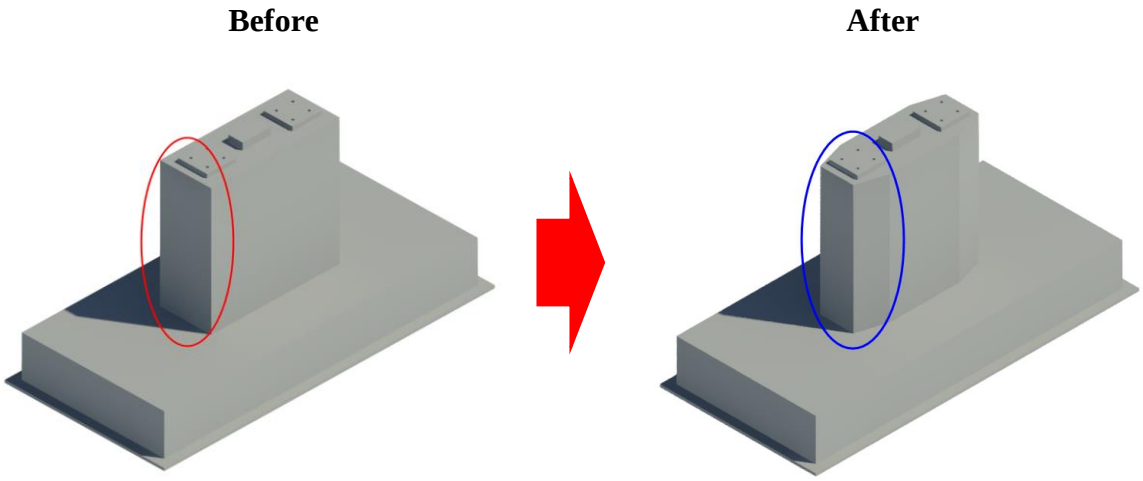


Figure 2. *The shape of the pier is changed in the 3D model.*

3D model of reinforced concrete structures of the entire project such as abutments,

piers, railing, ledge railing, deck slabs, bored cast-in-place pile are ideal for checking the arrangement or conflict of rebar bars and extracting concrete and steel volumes.

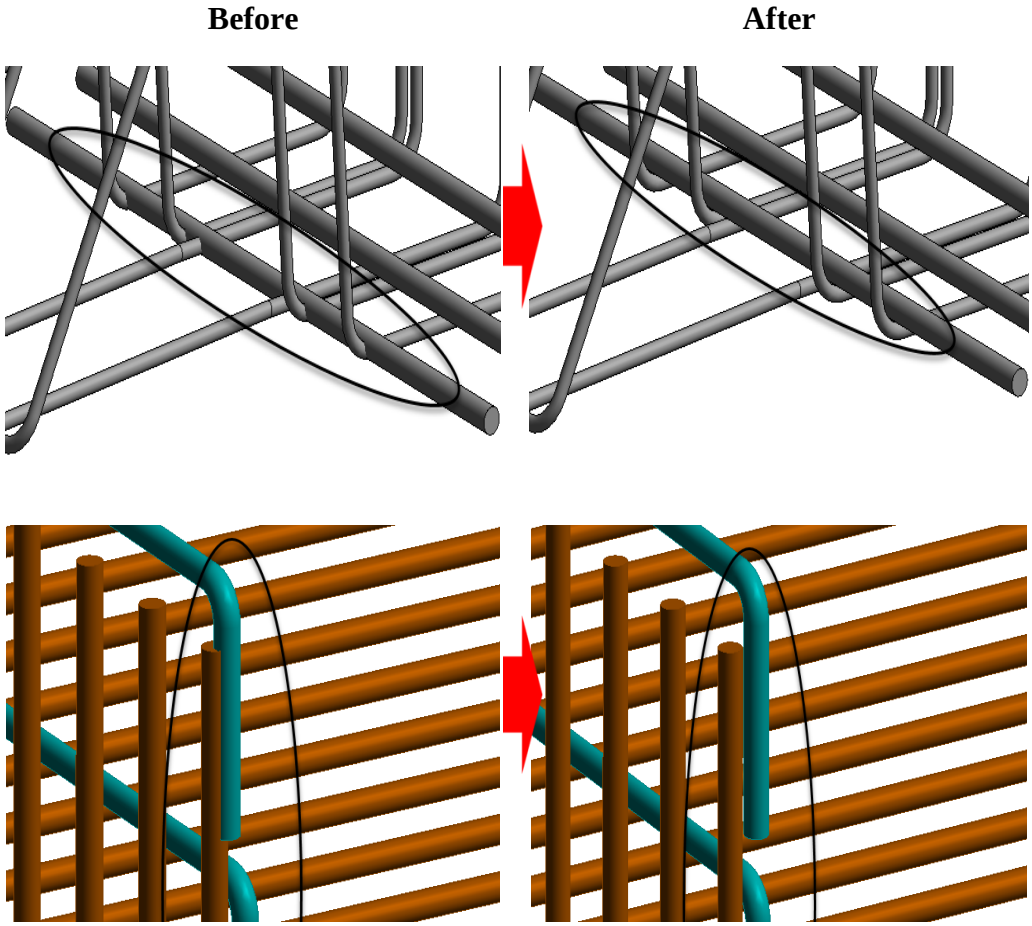


Figure 3. *Check and resolve conflicts of steel bars on the 3D model.*

Conflict detection helps to effectively check and report errors in the model, reducing risk of the design. For conventional design, errors are often discovered at construction step; but now, with conflict detection tools on BIM Model, they can be easily detected in the design step [3],[4]. This lead to save time, material and workforce for overcoming errors at construction.

2.2. The detailed design (LOD 350)

The BIM model accurately shows the level of clarity of information and details of the structure of host beams, abutments, piles, ... The number, size, shape, position, and direction of these members are showing clearly and linked with other systems [5]. Non-geometric information such as product specifications and specifications will be included in the model at the detailed construction design (LOD 350). The model provides enough information for exacting quantity, exporting the full 2D drawings and dividing phases for construction (Figure 4).

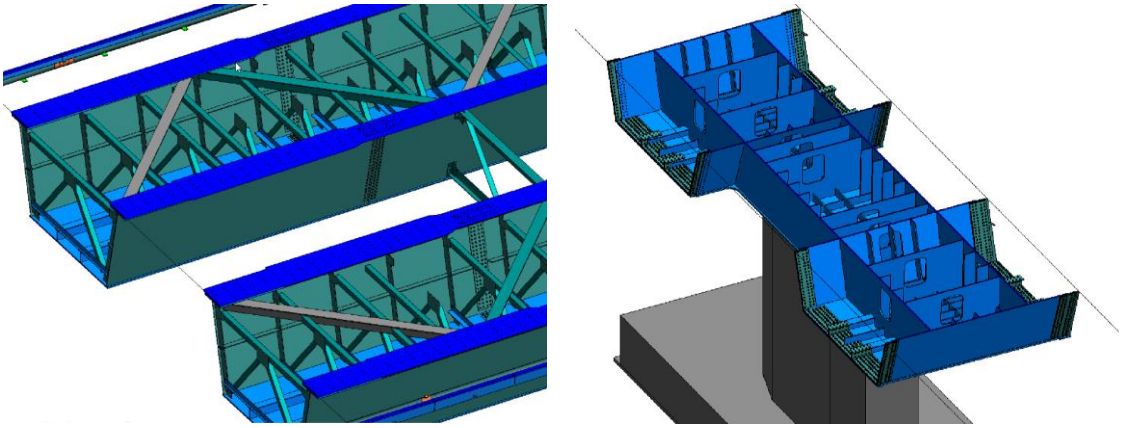


Figure 4. *The detailed of steel girders at LOD 350 step.*

For LOD 350, the detailed structure of steel girders includes ribs, anchor nails, and welding with assigned specifications (Figure 5), providing sufficient information for carrying out construction at site.

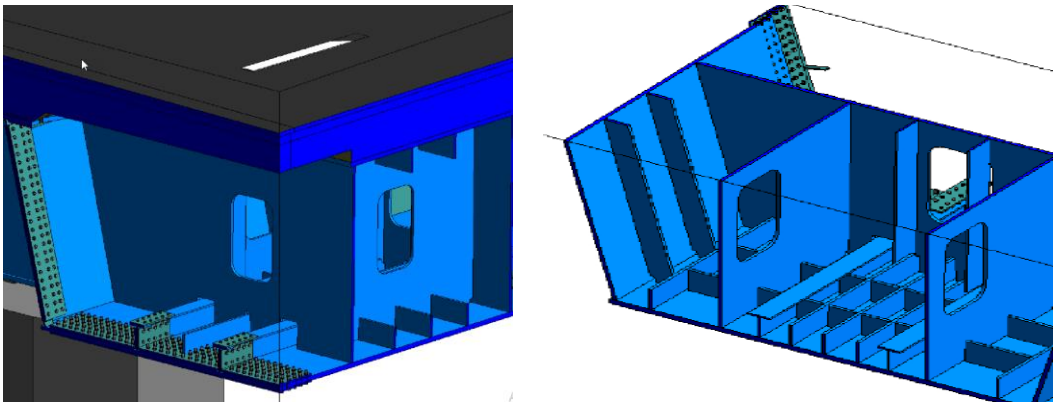


Figure 5. *The detailed of steel girder at pillar top.*

The abutment reinforcement arrangement is modelled using Dynamo programming and Revit software (Figure 6). Dynamo is a solution to improve the efficiency of 3D model in the design phase. The parametric design method using visual programming tools such as Dynamo helps maintain the original design ideas and improve the process of converting from idea to later design. In particular, the problems of creating complex geometry, sophisticated equipment and performing iterative operations have been solved. From there, the design options are evaluated correctly in a short time. Therefore, designing techniques using Dynamo will be beneficial in performing two tasks: customizing parameter geometry objects and communicating with external databases. During the design phase, the combination of Dynamo and Revit is appropriate, because Revit is considered a full database with geometric parameters [6], [7], [8].

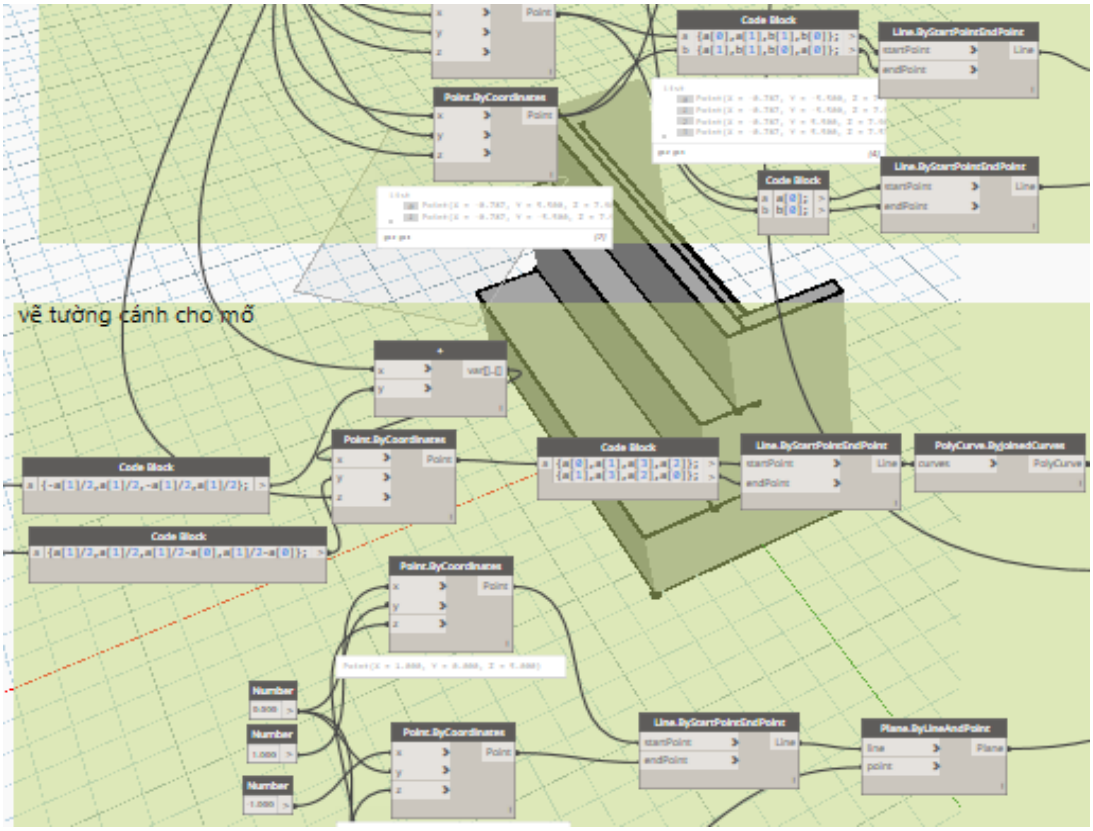


Figure 6. *The detailed abutment using the Dynamo program.*

In Figure 7, the drainage system on 550 – bridge is properly designed and arranged. An attribute data system, which needs to be considered for the operation and maintenance of bridge drainage systems such as drainage pipes, manholes, trash screens, is organized to build information, data on facilities basis and common application of BIM library models [9].

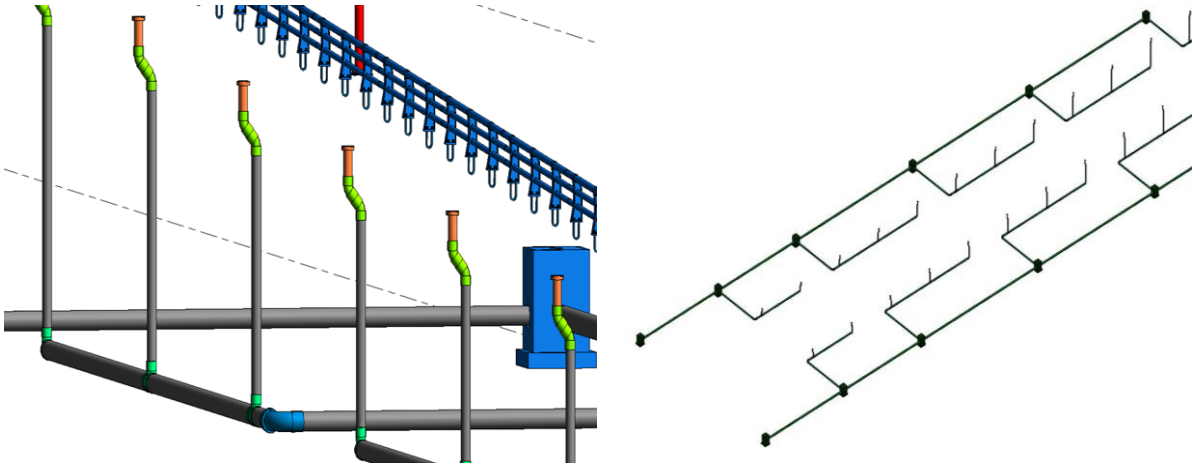


Figure 7. *Drainage design of 550 – bridge.*

Besides, some other components (bearings, panels, expansion joints, railings,...) are

designed in detail with technical parameters, non-geometric information and assigned to the model with sufficient information such as concrete grade, steel grade, size, ... to serve for material quantity and construction on site (Figure 8 and Figure 9).

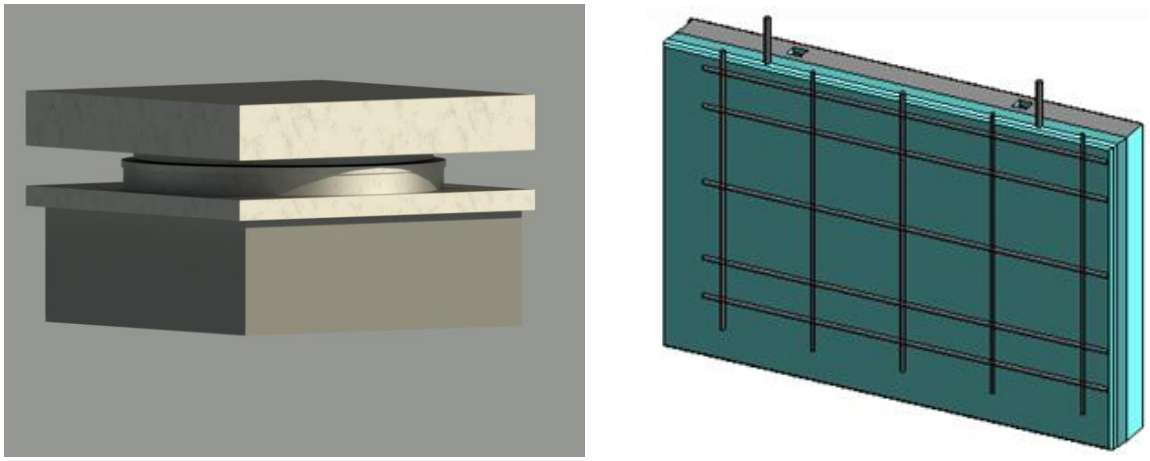


Figure 8. *The detailed of the bearing and panel retaining wall.*

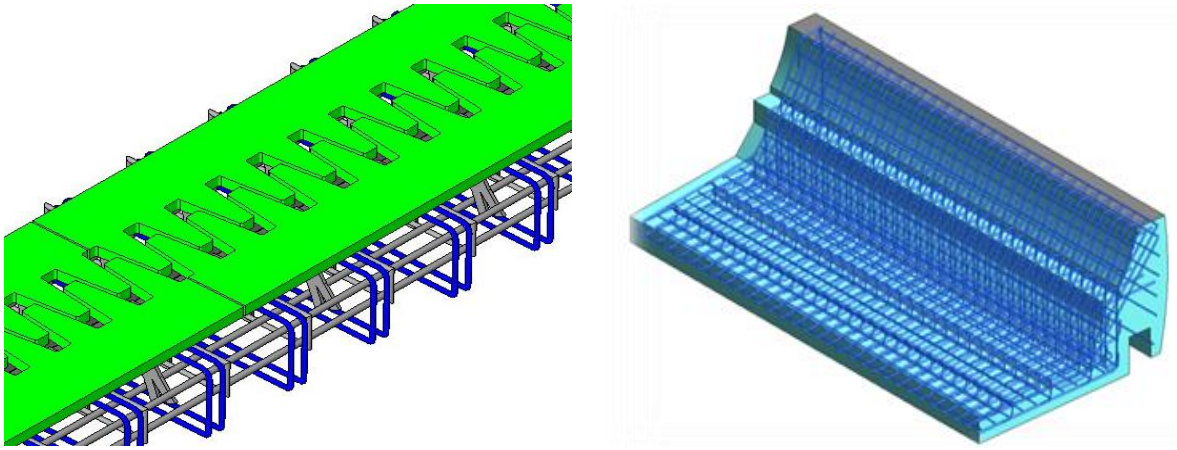
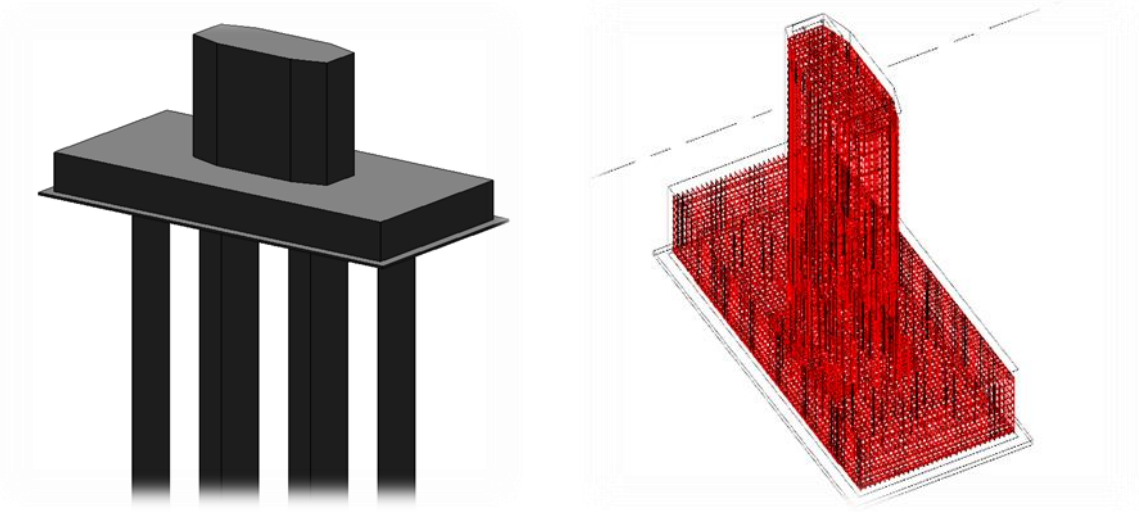


Figure 9. *The detailed of a flexible joint and ledge railing.*

2.3. Construction quantity

Material quantity is very important and an essential need when implementing any construction project. The project of bridge and crossroads 550 - Binh Duong includes many items such as abutments, pillars, beams, handrails, ... Construction work [10].

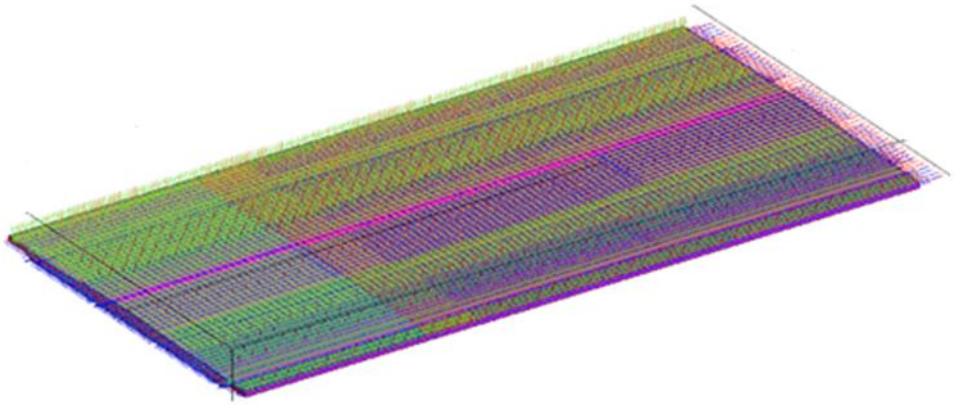
The abutment model is arranged in detail reinforcement and incorporates non-geometric information into the model, which makes it easy for material quantity work (Figure 10).



A	B	C	D	E	F	G
Partition	Bar Diameter	MINIMUM LENGTH	MAXIMUM LENGTH	Total Bar Length	Count	Mass
B1	32 mm	10850 mm	10850 mm	358050 mm	33	2259.35 kg
B1a	32 mm	7978 mm	8326 mm	282322 mm	34	1781.50 kg
B2	25 mm	4850 mm	4850 mm	354050 mm	73	1363.59 kg
B3	16 mm	5185 mm	5185 mm	82957 mm	16	130.87 kg
B4	16 mm	1549 mm	1549 mm	102218 mm	66	161.25 kg
B5	20 mm	10850 mm	10850 mm	358050 mm	33	882.56 kg
B6	20 mm	4850 mm	4850 mm	344350 mm	71	848.79 kg
B7	16 mm	10850 mm	10850 mm	173600 mm	16	273.86 kg
B8	16 mm	1549 mm	1549 mm	227650 mm	147	359.13 kg
B9	16 mm	3221 mm	3260 mm	351946 mm	108	555.21 kg
T1	32 mm	6469 mm	6684 mm	468598 mm	72	2956.92 kg
T2	16 mm	11134 mm	11134 mm	434208 mm	39	684.98 kg
T3	16 mm	4635 mm	4635 mm	111234 mm	24	175.48 kg
T4	16 mm	4384 mm	4384 mm	52608 mm	12	82.99 kg
T5a	25 mm	4567 mm	4567 mm	73069 mm	16	281.42 kg
T5b	25 mm	3831 mm	3831 mm	15326 mm	4	59.03 kg
T5c	25 mm	2831 mm	2831 mm	11326 mm	4	43.62 kg
T6a	20 mm	1611 mm	2143 mm	52554 mm	28	129.54 kg
T6b	20 mm	2144 mm	2144 mm	68605 mm	32	169.10 kg

Figure 10. *The detailed pier and export volume table.*

Application of BIM in measuring the volume of works has outstanding advantages such as being able to link two-way information exchange with other software such as Microsoft Excel, Project, ... If there is any change in the calculation Design changes, it will be automatically adjusted and the measurement volume will automatically update. Thus, errors in the process of determining and calculating material quantity will be reduced to a minimum. Since then, the time required for the measurement of quantity process will be reduced significantly [11],[12]. Figure 11 shows the detailed structure of the bridge deck, reinforced arrangement with assigned specifications and material table with detail of types, length, weight, and number.



A	B	C	D	E	F	G
Partition	Bar Diameter	MINIMUM LENGTH	MAXIMUM LENGTH	Total Bar Length	Count	Mass
B1a	14 mm	11700 mm	11700 mm	2433600 mm	208	2939.30 kg
B1b	14 mm	9815 mm	9815 mm	510380 mm	52	616.44 kg
B1c	14 mm	9255 mm	9255 mm	481260 mm	52	581.27 kg
B2a	22 mm	11700 mm	11702 mm	2386963 mm	204	7119.19 kg
B2b	22 mm	4733 mm	4733 mm	970265 mm	205	2893.84 kg
B3a	20 mm	1489 mm	1489 mm	11909 mm	8	29.35 kg
B3b	20 mm	1515 mm	6037 mm	2428255 mm	403	5985.41 kg
B4a	20 mm	3947 mm	3947 mm	793280 mm	201	1955.36 kg
B4b	20 mm	11654 mm	11654 mm	46617 mm	4	114.91 kg
B4c	20 mm	2865 mm	2865 mm	11462 mm	4	28.25 kg
B5a	14 mm	11700 mm	11700 mm	2281500 mm	195	2755.60 kg
B5b	14 mm	9815 mm	9815 mm	412230 mm	42	497.89 kg
B5c	14 mm	9255 mm	9255 mm	518280 mm	56	625.98 kg
B5d	14 mm	1515 mm	1522 mm	130768 mm	86	157.94 kg
B6a	18 mm	1439 mm	1439 mm	578417 mm	402	1154.85 kg
B6b	18 mm	1880 mm	1880 mm	755933 mm	402	1509.27 kg
B7a	12 mm	224 mm	280 mm	77304 mm	300	68.60 kg
B7b	12 mm	326 mm	326 mm	290508 mm	892	257.79 kg
B7c	12 mm	400 mm	400 mm	99121 mm	248	87.96 kg
B7d	12 mm	594 mm	600 mm	10113 mm	17	8.97 kg
G1	16 mm	678 mm	678 mm	277991 mm	410	438.54 kg
G2	16 mm	698 mm	698 mm	286235 mm	410	451.55 kg

Figure 11. *The detailed deck slab and export volume table.*

A	B	C	D	E	F
Family	Material: Area	Material: Name	Material: Volume	Mark	Count
BMC nhip biên	523 m²	Steel	0.00 m³	BMC	1
BMC nhip biên	487 m²	Concrete C35	161.58 m³	BMC	1
BMC giua tru va nhip	308 m²	Steel	0.00 m³	BMC	1
BMC giua tru va nhip	286 m²	Concrete C35	95.88 m³	BMC	1
BMC nhip biên	523 m²	Steel	0.00 m³	BMC	1
BMC nhip biên	487 m²	Concrete C35	161.58 m³	BMC	1
BMC giua tru va nhip	308 m²	Steel	0.00 m³	BMC	1
BMC giua tru va nhip	286 m²	Concrete C35	95.88 m³	BMC	1
BMC giua tru va nhip 2	304 m²	Steel	0.00 m³	BMC	1
BMC giua tru va nhip 2	290 m²	Concrete C35	95.88 m³	BMC	1
BMC giua tru va nhip 2	298 m²	Steel	0.00 m³	BMC	1
BMC giua tru va nhip 2	295 m²	Concrete C35	95.64 m³	BMC	1
BMC giua nhip1	376 m²	Steel	0.00 m³	BMC	1
BMC giua nhip1	349 m²	Concrete C35	115.14 m³	BMC	1
BMC giua nhip1	376 m²	Steel	0.00 m³	BMC	1
BMC giua nhip1	349 m²	Concrete C35	115.14 m³	BMC	1
BMC giua nhip2	377 m²	Steel	0.00 m³	BMC	1
BMC giua nhip2	350 m²	Concrete C35	115.46 m³	BMC	1

Figure 12. *Export volume table deck slab.*

2.4. 3D Simulation

The 550 - Binh Duong bridge is visually designed to effectively analyse images and imported directly into virtual reality software, without having to rebuild the model as conventional. With high resolution and special effects, the model is best for investors studying the project before construction. The ability of analyzing traffic flow in the construction area, especially during rush hour, helps to organize traffic effectively (Figure 13) [13],[14].

Figure 13. *Traffic simulation.*

3. CONCLUSION

BIM model of 550 – Binh Duong overpass for detailed construction design (LOD 350) will help engineers complete the design, simulation, and information management in all phases of the project. This project is also presented to prove the effective feasibility of applying BIM in transportation and helps to promote BIM adoption in complex bridge as well as infrastructure projects in the future [15].

ACKNOWLEDGMENT

We would like to thank the BROBIM Center, V7 Engineering & Technology CO.LTD, International conference on Applications of Artificial Intelligence in Transportation and University of Transport and Communications – Campus in Ho Chi Minh City for supporting this research.

REFERENCES

[1]. TS. Trần Hồng Mai, TS. Nguyễn Việt Hùng, TS. Tạ Ngọc Bình, Ths. Lê Thị Hoài Ân (2014), Tạp chí Kinh tế xây dựng, số 02.

[2]. Hyun Cho, Ki Hwan Lee, Song Hyon Lee, Ta Lee, Hyun Jun Cho, Sung Hoon Kim, Sang Hyeok Nam (2011), “Introduction of Construction Management Integrated System Using BIM

in the Honam High-Speed RailwayLot No. 4-2", The International Association for Automation and Robotics in Construction, 2011 Proceedings of the 28th ISARC, Seoul, Korea, tr. 1300 – 1305.

[3]. Wenping Liu, Hongling Guo, Heng Li and Yan Li (2014), "Using BIM to Improve the Design and Construction of Bridge Projects: A Case Study of a Long-span Steel-box Arch Bridge Project", Int J Adv Robot Syst, DOI:10.5772/58442.

[4]. Zita Sampaio (2015), "The Introduction of the BIM Concept in Civil Engineering Curriculum", International Journal of Engineering Education, tập 31, số 1 (B), tr. 305 – 305.

[5]. Quyết định 1057/QĐ-BXD Công bố Hướng dẫn tạm thời áp dụng Mô hình thông tin công trình (BIM) trong giai đoạn thí điểm, Bộ Xây Dựng ban hành ngày 11 tháng 10 năm 2017.

[6]. Nguyễn Mạnh Tuấn (2018), "Ứng dụng Dynamo cho dự án BIM trong giai đoạn thiết kế ý tưởng", Tạp chí khoa học công nghệ xây dựng, tập 12, số 1, tr. 71 – 76.

[7]. Dynamo BIM Community (2017), http://www.dynamobim.org/, truy cập ngày 20/02/2020.

[8]. San Francisco Dynamo User Group (2017), http://www.sfdug.org/, truy cập ngày 20/02/2020.

[9]. Jaehyun Park, Hyundong Lee và Pilljae Kwak (2015), "Case Study on the Application of BIM in Korea's Civil Engineering and Improvement of Interoperability of BIM Models", International Journal of Software Engineering and Its Applications, số 10, tr. 29 – 38.

[10]. K. Sears, G. Sears, and R. Clough, "Construction Project Management: A Practical Guide to Field Construction Management (5th Edition), "John Wiley and Son, Inc., Hoboken, New Jersey, 2008.

[11]. Hồ Văn Võ Sĩ, Hoàng Nhật Đức, Vũ Duy Thắng và Nguyễn Thị Bích Thủy (2016), "Ứng dụng mô hình thông tin xây dựng (BIM) vào việc đo bóc khối lượng công trình xây dựng", Tạp chí Khoa học và Công nghệ, tr. 68 – 74.

[12]. Yanran Wang, Gougang Wang, Jingru Wang (2017), "Research on the Application of BIM Technology in the Project Management of Hospital Construction", The Pacific Association of Quantity Surveyors, http://www.paqs.net/node/151, truy cập ngày 27/02/2020.

[13]. Ths. Huỳnh Xuân Tín, Đỗ Minh Truyền, Nguyễn Xuân Việt, Nguyễn Trường Giang, Bùi Vĩ và Võ Minh Khoa (2019), "Nghiên cứu lập mô hình thông tin công trình cầu cho bước thiết kế kết cấu nhịp cầu vượt ngã tư 550 – Bình Dương", Kỷ yếu nghiên cứu khoa học sinh viên năm học 2018 – 2019, Nhà xuất bản Giao thông Vận tải, Hà Nội, tr. 3 – 7.

[14]. Adam Strafaci (2008), "What does BIM mean for civil engineers? Road and highway projects can benefit from design using building information modeling", CE NEWS, tr. 62 – 65.

[15]. Nguyễn Việt Hùng (2015), "Nghiên cứu xây dựng lộ trình áp dụng mô hình thông tin công trình (BIM) nhằm nâng cao hiệu quả thiết kế, xây dựng và quản lý công trình tại Việt Nam", Báo cáo kết quả nghiên cứu đề tài khoa học, mã số RD 03 – 14, Bộ Xây Dựng.

DESIGN AND IMPLEMENTATION OF THE FUZZY-BASED PID CONTROLLER FOR THE HEATING SYSTEM

Trinh Luong Mien[1*], Vo Van An[2], Bui Thanh Tam[3]

[1]*University of Transport and Communications, No.3 Cau Giay Street, Lang Thuong Ward, Dong Da District, Hanoi, Vietnam.*

[2]*Binh Duong Economics and Technology University, No.530 Binh Duong avenue, Thu Dau Mot, Binh Duong, Vietnam.*

[3]*Thu Dau Mot University, No.6, Tran Van On street, Thu Dau Mot, Binh Duong, Vietnam.*

**Corresponding author: mientl@utc.edu.vn*

Abstract: Temperature control plays an important role in life, industry and transportation. In this paper, the author focuses on the design and implementation of the heating temperature control system, using a combination of the PID algorithm and fuzzy logic. Firstly, the authors identify the mathematical model of the heating system. Then, the authors design, simulate and evaluate the temperature control quality of the heating system on Matlab. Finally, the authors design the PIC16f control board with fuzzy based-PID algorithm to control the temperature of the heating system. Simulation and experimental results show that the proposed fuzzy-based PID control algorithm provides a good temperature control quality of the heating system when the system is disturbed.

Keywords: Fuzzy, PID, Heating system, Temperature control, PIC, microcontroller

1. INTRODUCTION

Heating systems are widely used in life, industry and transportation. The temperature control of heating systems has always attracted the attention of many researchers. The heating system is nonlinear object with the large time delay [1,3,6]. Moreover, the heating is always affected by noise when it is operated [13,16]. Therefore, determining the exact mathematical model of the heating system is extremely difficult [9,12]. For that reason, traditional PID control methods often do not really bring high quality [2,7,9,12]. Meanwhile, intelligent control methods, including fuzzy logic [8], have proved effective in many complex control applications [3,5,10,14].

Traditional control methods such as the state feedback control [2], PID [7,9,12] or the intelligent control method using the fuzzy controller [8] is not guaranteed the quality of the temperature control system when the system is affected by noise. The combination of the fuzzy logic and the PID controller has brought good temperature control quality to the heating system [10,11,15]. However, these control algorithms, implemented on microcontrollers, require the large computational volumes. These studies have also not mentioned elimination the noise problem.

The study of this paper focuses on the combination of fuzzy logic and PID method to

achieve better temperature control quality for the nonlinear, impact-noise heating system. Based on the physical heating oven, firstly the authors identify the mathematical model of the heating system. Then the authors propose the PID controller combined with fuzzy logic, in which PID's parameters are calculated based on fuzzy logic block. And then the design of PID controller, and fuzzy-based PID controller for the heating system are presented in this paper. After that, the authors simulate and assess the temperature control quality of the heating system on Matlab when the system is affected by noise. Finally, the authors design hardware using PIC16f microcontroller and develop the embedded software to install the fuzzy based-PID algorithm on PIC16f microcontroller for the temperature control of the heating system. The simulation and experimental results show that in all cases, the fuzzy based-PID controller always ensures the temperature control system operates stably with good quality.

2. IDENTIFICATION OF THE HEATING SYSTEM

To identify the heating system, the authors proactively applied voltage to the resistor and measured the output temperature of the heating system. The heating system has an input voltage signal *U(t)*, and an output temperature signal of *C(t)*. When we put the voltage is 100% of the heating system capacity, at the time of measurement we see that the temperature of the heating system has not changed immediately, it takes a certain amount of time to convert electricity into heat energy, heat transfer in the heating system and this time is the delay time of the heating system model. The temperature in the heating system will gradually increase to the maximum value, corresponding to the maximum power of the heating system.

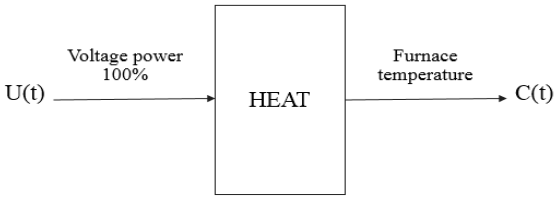


***Figure 1**. The identification process of the heating system.*

By this experimental method, the authors have determined the mathematical model of the heating system. The initial temperature is the ambient temperature 27^0C. The measured steady temperature value is 91.4^0C. The time to set the steady output temperature of the heating system is 30 minutes.

Time(m)	1	2	3	4	5	6	7	8	9
Temp(^{0}C)	27.8	29.1	30.5	32	33.5	35.9	38.5	41.3	44.7

10	11	12	13	14	15	16	17	18	19
47.8	51.2	54.9	58.2	61.6	65	68.2	71.1	74.2	76.8

20	21	22	23	24	25	26	27	28	29
79.4	81.7	84	86	87.7	88.9	89.9	90.5	91	91.4

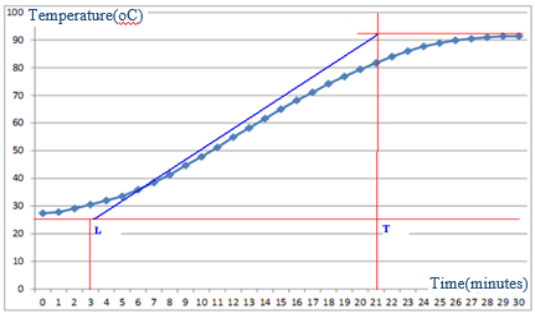


***Figure 2**. Measurement temperature curve of the thermistor heating system.*

The temperature curve of the heating system in the form *"S"*, so the mathematical model of the heating system is described as below [2]:

$$P(s) = \frac{k}{1+sT} e^{-Ls} \quad (1)$$

k: the object transfer coefficient, is the limit output value as time comes to infinity.

L: the time delay constant, period that the output signal does not respond immediately.

T: the inertial time constant

Based on the temperature characteristic curve of the heating system, the authors draw the tangent line with the inflection point of the curve *"S"*, the parameters of the heating system as follows:

$$k = \frac{91.4 - 27.8}{100} = 0.64 \quad (2)$$

Besides, it is easy to identify *L=225 second* and *T=1230 second*

Therefore, we obtain the transfer function of the heating system as follow

$$P(s) = \frac{0.64e^{-225s}}{1230s+1} \quad (3)$$

It is easy to see that the heating system model is nonlinear.

For the design of the PID controller, the authors apply approximately the non-linear transfer function of the linear transfer function, as follow

$$P(s) = \frac{0.64}{276750s^2 + 1455s + 1} = \frac{k}{(1+T_1 s)(1+T_2 s)} \quad (4)$$

So, we have: $k=0.64;\ T_1=225;\ T_2=1230.$

3. CONTROLLER DESIGN AND SIMULATION FOR THE HEATING SYSTEM

3.1. Design of the PID controller for the heating system

According to Ziegler Nichols-1 (ZN1), the PID controller parameter for the resistance furnace, is determined in the table below [2].

$$G_{PID}(s) = K_P(1 + \frac{1}{T_1 s} + T_D s) = K_P + \frac{K_I}{s} + K_D s \quad (5)$$

	K_P	T_I	T_D
PID	$1{,}2.\frac{T_2}{T_1.K}$	$2.T_1$	$0{,}5.T_1$

Therefore, the authors calculated:

$K_P = 10{,}256;$ $T_I = 450$ ($K_I = 0{,}0227$) $T_D = 112{,}5$ ($K_D = 1153{,}8$)

In order to limit the impact of noise on the system, the authors design the PID controller

by Chien-Hrones-Reswick method (CHR) [2].

	K_P	T_I	T_D
PID	$\frac{19.b}{20.a.K}$	$\frac{12.a}{5}$	$\frac{21.a}{5}$

Therefore, k=0.64; a= T_1=225; b=T_2=1230

Since then, the authors determine the PID coefficient by CHR as follows: K_P = 8,114; K_I = 0,015; K_D = 7667,73

After adjusting the parameter of PID controller by CHR1 with K_P = 8,114; K_I = 0,015; K_D = 1500 the authors find the response of the system having better quality.

3.2. Design of the fuzzy based-PID controller for the heating system

Fuzzy based-PID controller is synthesized based on the structure of PID with K_P, T_I and T_D parameters determined according to the corresponding fuzzy stages: fuzzy stage of calculating K_P, fuzzy stage of calculating T_I, fuzzy stage of calculating T_D. These fuzzy stages have two inputs as error $e=T_d^o-T^o$ and error rate de/dt. So, the parameters of PID are automatically adjusted continuously in a specified range, thus improving the control quality for the temperature control system [1].

Each fuzzy stage has two inputs and one output: First input $e = T_d^o-T^o$ (Error between the set temperature and the actual temperature in the heating system), The second input is the temperature error rate $\frac{de}{dt} = \dot{e}$, The output corresponding to the fuzzy stage of calculating K_P is the K_P ratio, with the fuzzy stage of calculating T_I is the T_I integration coefficient, with the fuzzy stage of calculating T_D is the K_D differential coefficient.

The fuzzy calculation stages K_P, T_I, T_D have two inputs e and $\dot{e}$, one output K_X, in which K_X can be K_P, or T_I, or T_D is presented as shown below.

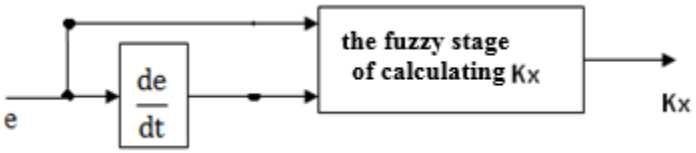


***Figure 3**. The general structure of the fuzzy stage calculating the PID parameters.*

The fuzzy stage of calculating of PID coefficients is designed on Matlab according to Mandani with two inputs e and $\dot{e}$, one corresponding output for the three fuzzy stages of calculating K_P, T_I and T_D.

The output of fuzzy stage K_P, denoted by P selected in the value domain [0, 20].

The output of fuzzy stage T_I, denoted by I selected in the value domain [0, 0.05].

The output of fuzzy stage T_D, denoted by D selected in the value domain [0, 2000].

For K_P, T_I the signal error e, denoted by ET is selected in the value range [-20, 20]; The error rate $\dot{e}$, denoted by DET is selected in the value range [-2, 2].

For T_D the signal error e, denoted by ET is selected in the value range [-100, 100]; The

error rate ė, denoted by DET is selected in the value range [-10, 10].

Perform fuzzy input variables as follows: Error of ET={HQ,HD,HI,HV,HL}; Derivative error DET={TA,TZ,TI,TV,TL}.

Fuzzy output variables K_P, T_I and T_D corresponding to the language variables P, I, D we have: P={VD,VG,VI,VV,VL}; I={VD,VG,VI,VV,VL}; D={VD,VG,VI,VV,VL}.

Select the trapezoidal membership function for input and output linguistic variables.

With the number of fuzzy sets of inputs 5 and here using 2 inputs for each fuzzy stage calculating K_P, T_I and T_D we have the total of 5x5=25 fuzzy rules for each fuzzy stage. Based on the heating system specification and the principle of PID parameter adjustment to control the thermistor heating system, the authors built the fuzzy rules table for the fuzzy stages of calculating K_P, T_I and T_D as shown:

Table 2: *Fuzzy rules for K_P, T_I and T_D.*

		ET				
		HQ	**HD**	**HI**	**HV**	**HL**
DET	**TA**	VD	VD	VD	VD	VD
	TZ	VD	VG	VD	VD	VD
	TI	VD	VD	VI	VI	VV
	TV	VD	VD	VI	VV	VL
	TL	VD	VD	VV	VL	VL

Building the simulation diagrams of the heating system temperature control system using fuzzy based-PID controller with parameters K_P, T_I and T_D, determined through the fuzzy stages on Matlab as follows.

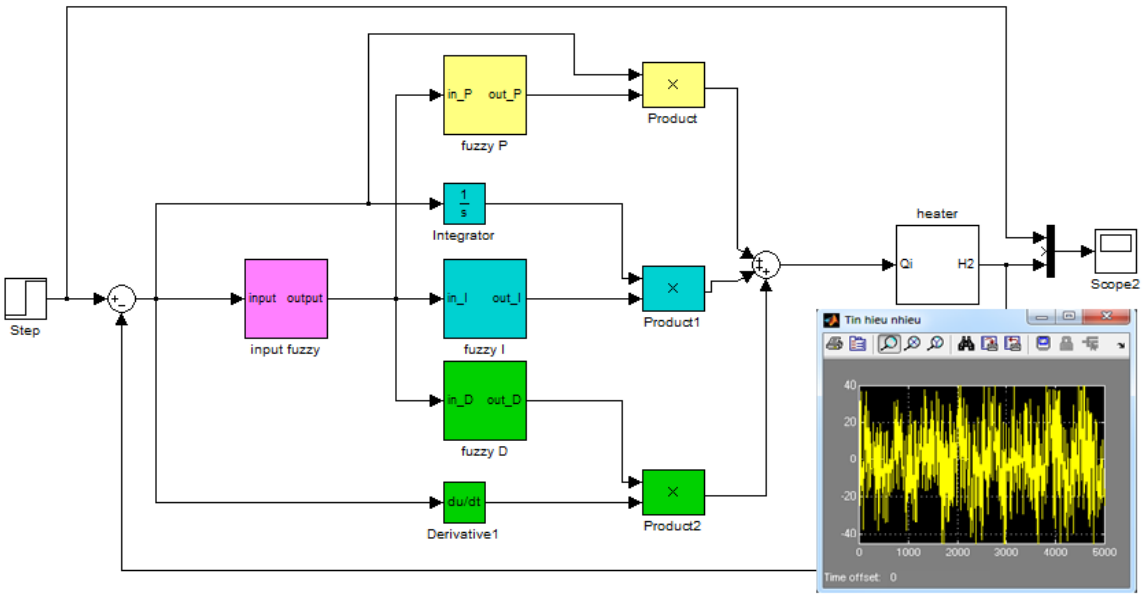


Figure 4. *The heating temperature control simulation using the fuzzy based-PID controller and the sine-noise signal impacted the system.*

3.3. Simulation and quality evalution of the heating temperature control system

The simulation of the heating temperature control system is done with three designed controllers above: PID-ZN1, PID-CHR and fuzzy based-PID.

When there is not noise, the response of the system corresponding to three controllers as shown Fig.5. The overshoot of the control system with fuzzy based-PID controller (PID-FZ) is better than conventional PID controller (PID-ZN1 or PID-CHR). Specifically, with the fuzzy based-PID system with no overshoot, PID-ZN1 or PID-CHR an overshoot of about 25%, the steady time of the system with the fuzzy based-PID controller is very short (nearly 20 seconds), the fuzzy based-PID controller for quick response. Meanwhile, the steady time of the PID-ZN1 or PID-CHR is greater, about 150 seconds. The steady-state error of the system with the fuzzy based-PID controller is small or none, ensuring global system.

So that fuzzy based-PID controller makes the system quality control superior to the use of conventional PID controller (PID-ZN1 or PID-CHR).

When there is sine-noise signal: In fact, the environment is not ideal. Controllers are always affected by external noise, so that it makes to change the actual signals. the response of the system corresponding to three controllers as shown Fig.6.

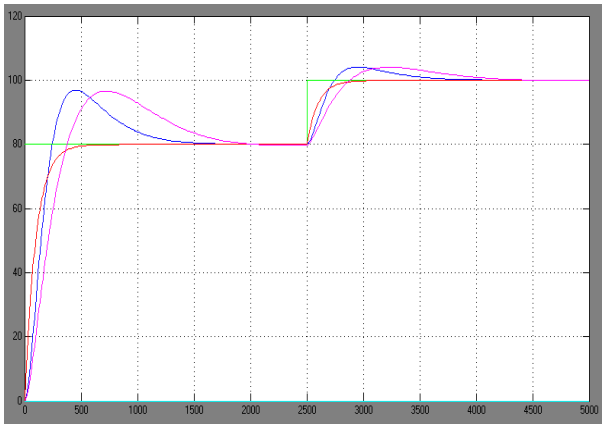


Figure 5. *Response without noise*

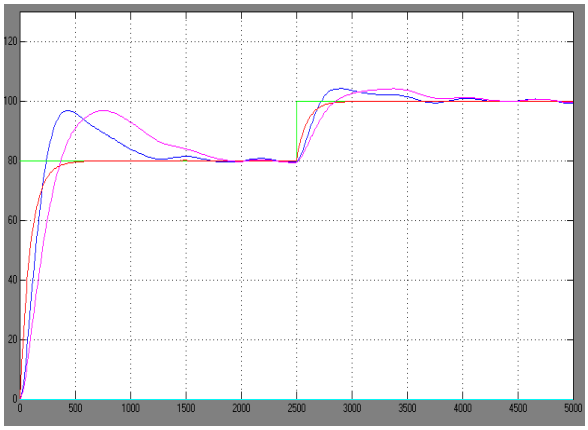


Figure 6. *Response with the sine-noise*

Based on the response of the system to each controller in Fig.6, the authors have some conclusions as follows: use of conventional PID controller (PID-ZN1 or PID-CHR) makes the system unstable. As for the fuzzy based-PID controller, the system quality still ensures the system operates stably.

4. DESIGN OF THE PIC16F CONTROL MODULE FOR THE HEATING SYSTEM

4.1. The block diagram of the PIC16f control module

In this section, the authors present the design of the PIC16F control module for the heating system. This PIC16f control module is designed by using PIC16f microcontroller [4].

The PIC16F control module block diagram for the heating system is shown Fig.7.

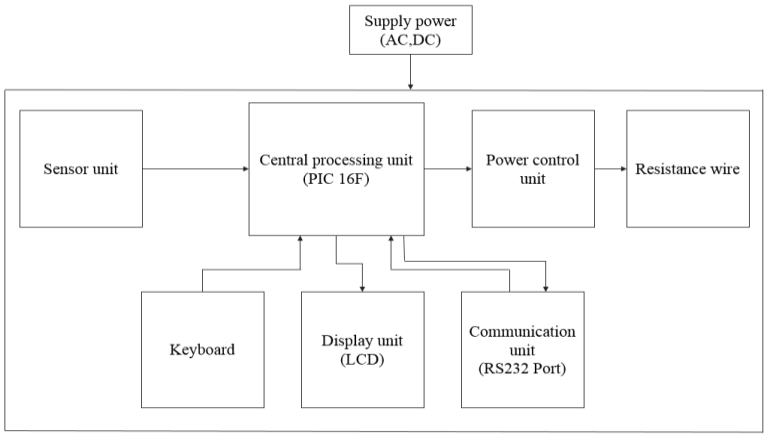


Figure 7. *The block diagram of the PIC16f control module for the heating system.*

The supply power is responsible for supplying power to the entire system; The control object here is the thermistor wire wrapped around the copper pipe causing the heat in the copper pipe to rise, it is controlled via the voltage signal of the MOSFET, the temperature in the copper pipe is measured by a temperature sensor; Display unit is responsible for displaying heating system temperature and set temperature; The keyboard is responsible for entering the set-point temperature; The central processing unit converts from analogue signal to digital signal and performs the functions of fuzzification, fuzzy-rules, defuzzification and generating signals to control the power unit to stabilize the temperature in the heating system; The power control unit has the task of changing, controlling the power supply of the thermistor. The RS232 port unit is responsible for communicating with the computer.

4.2. Schematic and layout circuit diagram of the PIC16F control module

Schematic circuit diagram of the PIC16f control module for the heating system includes the schematic circuit diagram of the PIC16f central processing unit, LCD display block; the power block, power control unit, RS232 communication block.

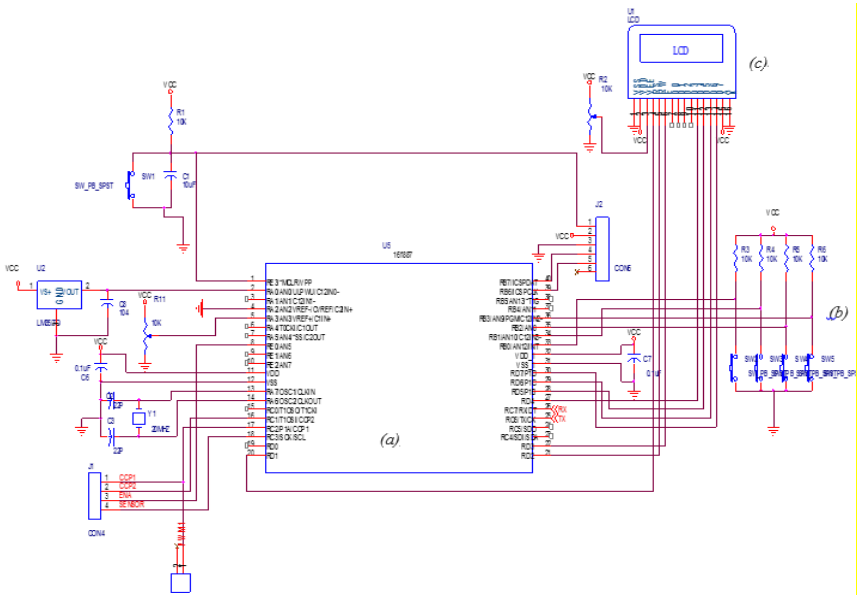


Figure 8. *Schematic diagram of control unit (a), keyboard (b), LCD unit (c).*

Based on the schematic circuit, the layout circuit diagram as shown below:

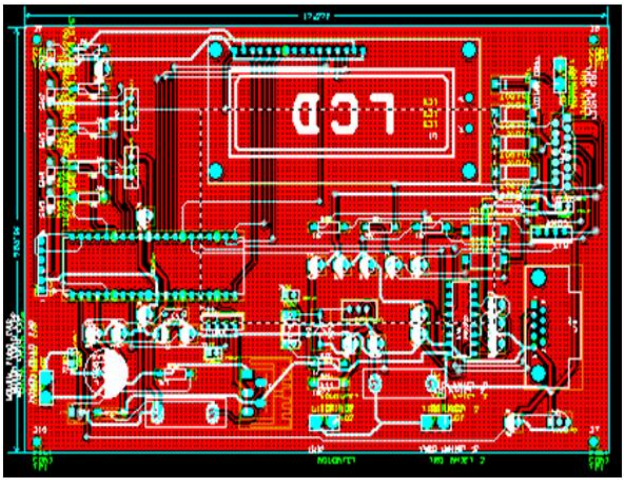


***Figure 9**. Layout circuit diagram.*

4.3. The heating temperature control algorithm on PIC16f microcontroller

The heating temperature control algorithm is installed on the PIC16f module as below.

In the figure 10, the Fuzzy-PID subroutine block is built based on the above design of the fuzzy based-PID controller in part 3.2.

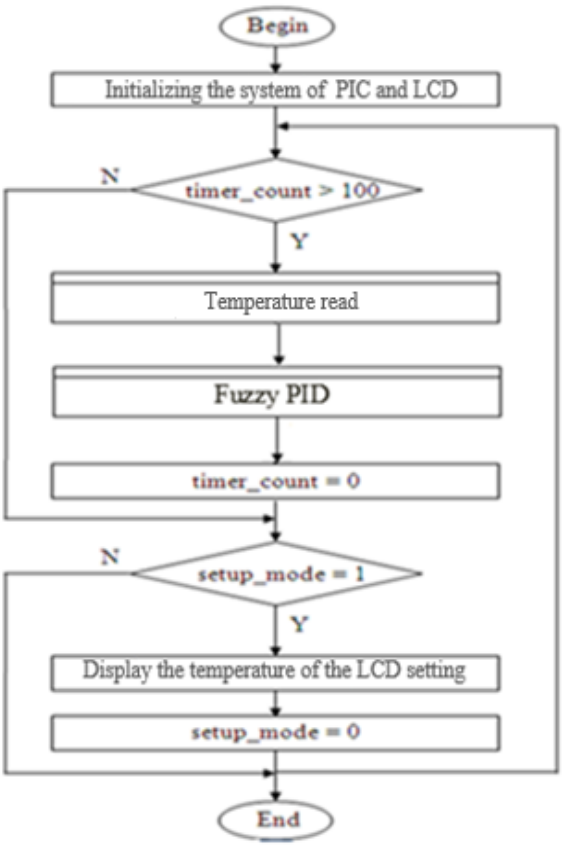


***Figure 10**. The main program flowchart on PIC16f with the fuzzy based-PID controller.*

4.4. Experimental control for the heating system in the laboratory

Experimental implementation for controlling the heating temperature in the laboratory using PIC16F control module, installed the fuzzy based-PID algorithm as Fig.12.

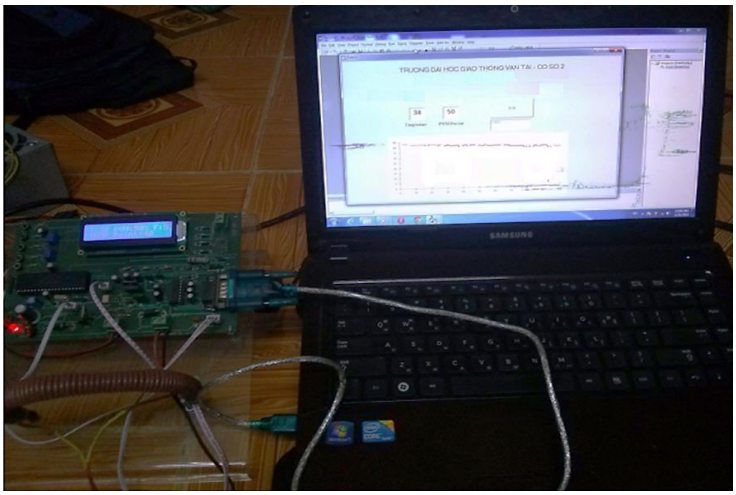


Figure 11. *Experimental heating temperature control system using PIC16F with the fuzzy based-PID.*

On the monitoring control interface, developed on VC#, the temperature characteristic curve of the heating system as shown below.

At the beginning of the test, the PIC16F microcontroller with fuzzy based-PID algorithm, will output the PWM signal to create a voltage for the resistor, after the delay time, the temperature in the heating system changes from 0^{o}C up to 36^{o}C, then from 36^{o}C up to 39^{o}C and maintain at a set temperature of 40^{o}C. When the oven temperature exceeds 40^{o}C, the PIC16f control module interrupts the PWM signal so that the oven temperature rises down to the set value.

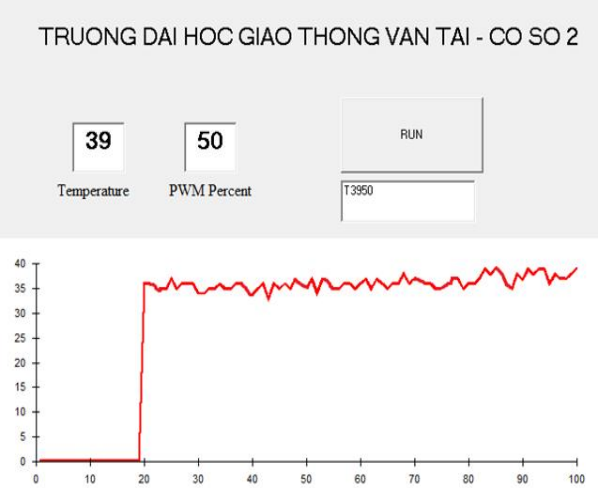


Figure 12. *The temperature characteristic of the heating system with T_d=40^{o}C .*

4. CONCLUSION

The paper presented a new detail approach of the heating temperature control system: from the heating system identification using experimental method, after that design of the controller for the heating system using the traditional PID (Ziegler-Nichols-1 method, Chien-Hrones-Reswick method) and propose the fuzzy based-PID controller based on combining fuzzy logic with PID controller, and then the simulation and the quality evaluation of the heating temperature control system with impact noise are performed on Matlab. Finally, the authors design and implement the PIC16f control module for the heating system to test the proposed fuzzy based-PID algorithm in this paper. Experimental results in the laboratory show that the PIC16f control module is installed the fuzzy based-PID algorithm for better control quality, as comparing to traditional PID. Especially, it has good anti-interference ability. The monitoring-control program on the computer also allows effectively the data management and manipulation of the system.

Hopefully with this research result, the PIC16f control module is able to be continued researching to apply in practical application.

REFERENCES

[1]. Phan Xuan Minh, Nguyen Doan Phuoc, Fuzzy control theory, Hanoi Science and Tech. publishing, 2004.

[2]. Aidan O'Dwyer, Handbook of PI and PID Controller Tuning Rules, Imperial College Press, 2003.

[3]. L. X. Wang, A course in Fuzzy Systems and Control, Prentice Hall International, Inc.

[4]. Microchip Technology Incorporated, PIC 16F Family Data Sheet, DS4281D, 2007.

[5]. Nguyen Chi Ngon, Fuzzy PI controller: From design to application, Science magazine, Can Tho Univ., 2011.

[6]. Moh Syahrir Bin Ribuan, Development Of Fuzzy Logic: Temperature Controller Implementation – Microcontroller Based, University Teknikal Malaysia Melaka, 2009.

[7]. Nguyen Thi Thanh Quyh, Pham Van Them, Design of PID controller and FLC controller (Fuzzy Logic Control) for heating system, Journal of science and technology.

[8]. P. Singhala, D. N. Shah, B. Patel, Temperature Control using Fuzzy Logic, International Journal of Instrumentation and Control Systems, 2014.

[9]. Yu Feng Zhang, Ming Li and Jing Min Dai, PID Heating and Temperature Control Method Based On Dynamic Assignment, 5th International Conf. on Mechatronics, Materials, Chemistry and Computer Eng., 2017.

[10]. SHI dequan, GAO guili, GAO Zhiwei, XIAO Peng, Application of Expert Fuzzy PID Method for Temperature Control of Heating Furnace, Sciverse sciencediret, 2012.

[11]. Sneha S Patole1, Shailendra K Mittal, Fuzzy PID Controller Design for Heating Control System, International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, 2007.

[12]. V.Kabila, Glan Devadhas, Comparative Analysis of PID and Fuzzy PID Controller

Performance for Continuous Stirred Tank Heating system, Indian Journal of Science and Technology, 2015.

[13]. Qingjie Yang, Guohou Li, Xusheng Kang, Application of fuzzy PID control in the heating system, Chinese Control and Decision Conference, 2008.

[14]. Jiangjiang Wang, Dawei An, Chengzhi Lou, Application of Fuzzy based-PID Controller in Heating Ventilating and Air-Conditioning System, International Conf. on Mechatronics and Automation, 2006.

[15]. L. D. Vijay Anand, Lakshmi Asok,, Design of Fuzzy PID with Expert Control for a Temperature Process, International Journal of Engineering Research & Technology, 2014.

[16]. Gao Guofang,Li Zhengzhong, Design of Heating Furnace Temperature Control System Based on Fuzzy based-PID Controller, International Conf. on Information Engineering and Applications, 2012.

SMART LIBRARY AND INFORMATION RESOURCES SHARING FOR SMART CITY

Vuong Tien Do[1*], Hieu Tran Nguyen[1]

[1]*University of Transport and Communications, No 3 Cau Giay Street, Hanoi, Vietnam.*
[*]*Corresponding author: dtvuong189@gmail.com*

Abstract: Can the smart city provide a new perspective for public and academic libraries? How does the smart city impact the libraries as cultural and scientific assets? And how can libraries contribute to the development of the smart city? This article introduces and defines basis smart library conceps and presents a domain model of smart library. The domain model is used to introduce to relace the implementing in the context of smart city can be described in three solutions, i.e., building big data management software 4.0, technology for smart library, smart services.

Keywords: smart library, 4.0 library, smart library model, information resources sharing, smart city.

1. INTRODUCTION

In Industrial Revolution 4.0, many modern technologies allow the real world (people, society, cities, homes, transportation, cars, airplanes, etc.) connect to the virtual world (digital space, electronic devices, computers, mobile devices, etc.) making everything connected smarter. They are also applied to create the smart cities, smart libraries, smart factories,…

A smart city is an urban area that uses different types of electronic Internet of Things (IoT) sensors to collect data and then use insights gained from that data to manage assets, resources and services efficiently [16]. This includes data collected from citizens, devices, and assets that is processed and analyzed to monitor and manage traffic and transportation systems, power plants, utilities, water supply networks, waste management, crime detection, information systems, schools, libraries, hospitals, and other community services [10, 17].

A smart library is developed on the basis of modern digital technologies such as artificial intelligence (AI), big data, cloud computing, Internet of things (IoT) ... providing friendly, smartly and quckily library services for users (in both physical and digital spaces). Strongly supported by the artificial intelligence, users can interact with the smart library like communicating with a real human being [18]. The smart library generations based on developing of Internet and IT through stages 1.0, 2.0, 3.0, 4.0 in both Physical Space (Traditional Library) and Digital Space (Digital Library) is summarized in the following table [4] :

Table 1: *Smart library generations [2,5].*

Generations	Characteristics	Library application	Technology platform
Library 1.0 (1995-2005)	Information connection	Library information portal; Search by keywords; Tree of knowledge; Email…	HTTP, Client/ Server, HTML, Java, Flash, ...
Library 2.0 (2005-2010)	People connection	Library blog; Facebook library; Youtube library; Identify keywords by users; Search based on social network user behavior; Cloud storage services ...	AJAX, SOAP, RSS, SaaS, PaaS, IaaS, ...
Library 3.0 (2010-2015)	Knowledge connection	Ontological language; Database semantic; Semantic search; Search in natural language; Knowledge Database, Knowledge Map ...	RDF, XML, OW, SPARQL, SWRL, ...
Library 4.0 (2015-2025)	Everything / smart connection	Smart physical space: biometric security identification; intelligent reader management; smart bookshelves; automatic loan repayment; robot librarian; smart reading room ... Smart digital space: virtual assistant (digital librarian); smart search; reading guide - smart study ...	AI, IOT, Big Data, Robotics, Quantum Computing, Blockchain…

In the smart library space, users will communicate with robot librarians or virtual library assistants, their interactions will be through natural language (voice, writing). They not only search and access the library's large data sources, but also use IoT-connected library facilities. Physical space and digital space will be very friendly, easy and smart when communicating with private users and providing endless sources of information to private users [17]. Therefore, training and equipping library knowledge and information knowledge for private users is extremely important, such as determining the purpose and needs of information, using information search tools and assessments. and use the information found effectively ... It is these knowledge and skills that make them users - smart readers of the 4th generation smart library; helping them not to be submerged in the big data universe but also help them to smartly identify directions, search and use information effectively, save time and effort; help them promote a culture of reading, self-study and lifelong self-study, love and discovery - knowledge creation, etc. [2, 7, 11]

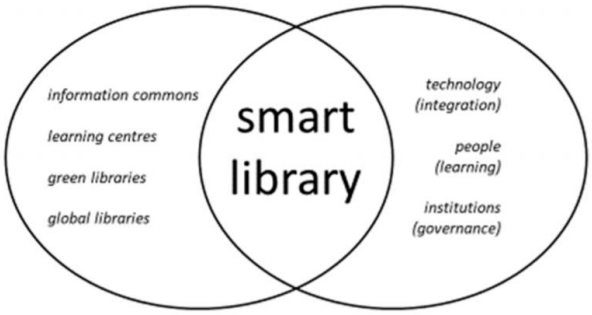


Figure 1. *The smart library at the interface between library concepts and the smart city [6].*

2. PROPOSAL MODEL FOR SMART LIBRARY AND INFORMATION RESOURCES SHARING FOR SMART CITIES

The smart city has been described as an "urban labeling" phenomenon, a "fuzzy concept", rather than a consistent framing or one-size-fits-all definition [13]. In the same way, other authors have characterized the smart city as an ambiguous, generic, and optimistic concept for the "city of the future", intended chiefly as an "efficient, technologically advanced, green and socially inclusive city" [14].

Several meanings and connotations have been linked to the term of "smart", e.g., efficient, sustainable, equitable, livable, instrumented, interconnected, intelligent city: "The use of Smart Computing technologies to make the critical infrastructure components and services of a city more intelligent, interconnected, and efficient. A city that gives inspiration, shares cultures, knowledge, and life, a city that motivates its inhabitants to create and flourish in their own lives" [13]. The relevant point for the development of public and academic libraries is not this assemblage of "pre-existing urban imaginaries" [14]. Two other aspects are of interest, i.e. the focus on information and the multidimensionality of the concept [9].

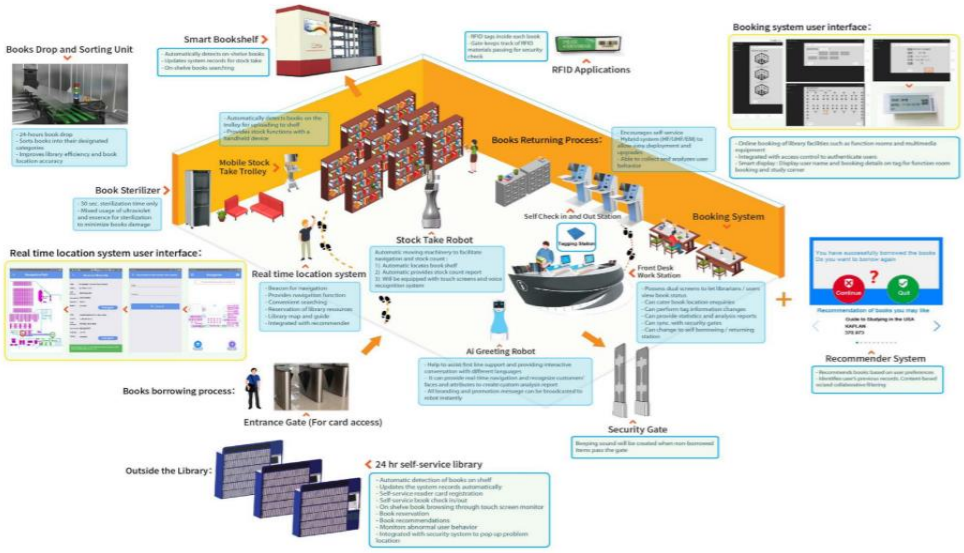


Figure 2. *Smart Library functions.*

Proposal model smart library

Based on the results of research above, the author boldly proposes his tissue to develope libraries 4.0 in universities over the world. The common models contain three main components as Data - Technology - Service and it also has the influence, the impact of the IoT, AI, Big data, Robots, cloud computing and interactive tools, smart devices such as IoT, RFID devices, smartphones and others [1].

Technology is one of the most elements in building the smart libraries. Huge data sources, extensive queries and repetition lead to library information service systems being able to analyze, "talk", "exchange" and "discuss" with scholars like a colleague. 100% of documents have RFID chips, mainly accessing libraries via mobile devices. Search document information via voice command for supported software. The web 4.0 features are also thoroughly applied by Library 4.0 [16, 17].

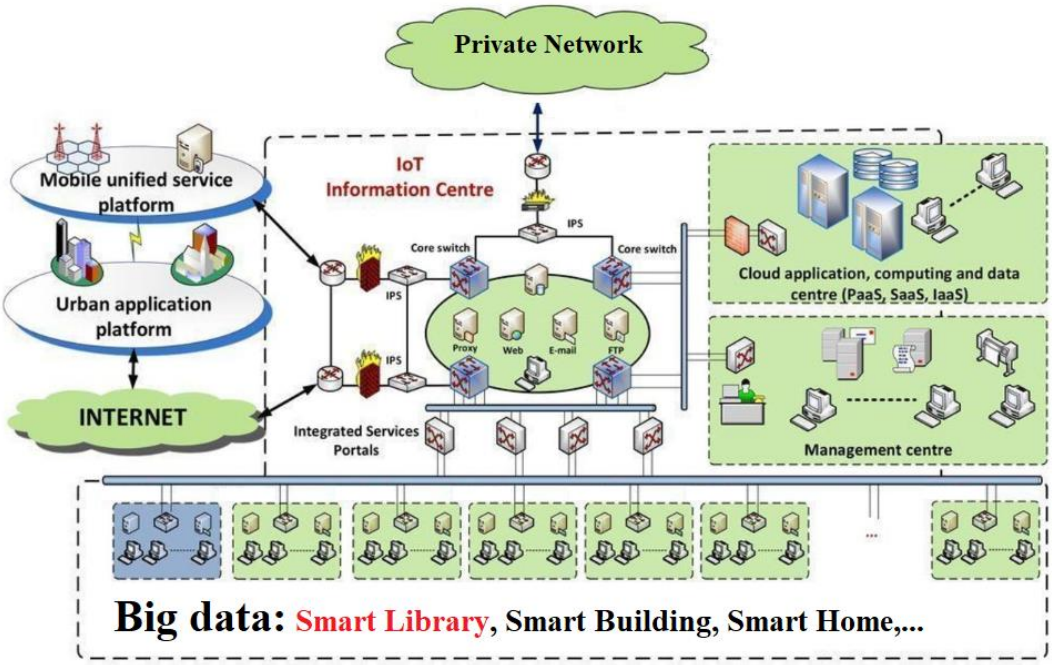


Figure 3. *Model of smart library and information resource sharing for smart city.*

The library's resources are mainly digital information and library trends connected to Twitter, LinkedIn, Facebook,... are common. The libraries of the same group/ field of training, research and service subjects will work together to maximize the effectiveness of information resources and services. Identified users via face recognition, voice, fingerprint applications are widely applied. Users can freely go to libraries regardless of their subjects and education levels. The physical library accepts the features of library 4.0 to change storage and academic spaces into space for collaboration and creativity. The keywords for library 4.0 will be smart library, big data, research environment in parallel with practice, open source, cloud computing, ... [1,8]

Smart library model implementation solutions

In order to develop and apply technology 4.0 for smart cities in Vietnam in the near future, the authors propose the following groups of solutions:

Firstly, building big data management software 4.0:

The current Vietnamese libraries have continuously digitized the building of endogenous resources; collections of dissertations, theses, scientific research results, books,... create very large data sources. However, they are located independently and distributed in libraries (mainly managed by the open source Dspace software). Therefore, to be able to connect all of the above data stores together to form a shared digital library system, unify a catalog system of metadata that points to digital resources in distributed databases. This canopy, just a single search command can query all the digital resources stores of the above libraries is a difficult but completely solved problem when we apply search technology 4.0 [2].

In the proposed model, the branch of the data represents four main elements: printed documents, endogenous digital documents, open access documents and electronic databases. These are also the 4 main types of documents that university libraries currently have and focus on. The shift from the type of printed materials to digital documents and its convenient and effective use made the investment balance of university libraries gradually shift to the electronic data. is ebook and ejournal. According to recent research data: MIT Library shows that the use of electronic documents by readers increases every year from 21 % to 35 %, Cambridge Univesity Library from 32 % to 48 %, Harvard Univesity Library from 25 % to 38 % so the investment in several hundred electronic databases includes millions of e-books and electronic journals compared to some electronic databases of Vietnamese university libraries are too big a difference. With a large number of Open Access academic journals published nowasday, scientific topics cover the field, Vietnamese university libraries can also take advantage to introduce to users the ways to effectively use this information resource [1, 3].

Secondly, technology for smart library:

The methodology for designing the smart library system development is based on two components; which are: the system hardware architecture and the software development.

System Architecture

The platform architecture consists of the technology that supports the IoT structure and the hardware platform. The technologies used in the IoT implantation are: The Internet, IPv6, RFID, Wireless Sensor Networks (WSNs) and the cloud system. The IoT employs a set of technologies in order to provide connectivity between sensor devices. Connectivity is a crucial issue that will be provided by the Internet and wireless technologies. The physical objects, actuators and controllers will be connected together through the Internet connection, with a unique identification.

These smart technologies facilitate the IoT implementation. WSNs helps establishing the non-centralized network between nodes and sending data to the gateway. RFID provides unique identification of the objects and enables to track them. Cloud servers are available in two types: Of private and public schemes to store and administrate all information [7].

Platform scenario

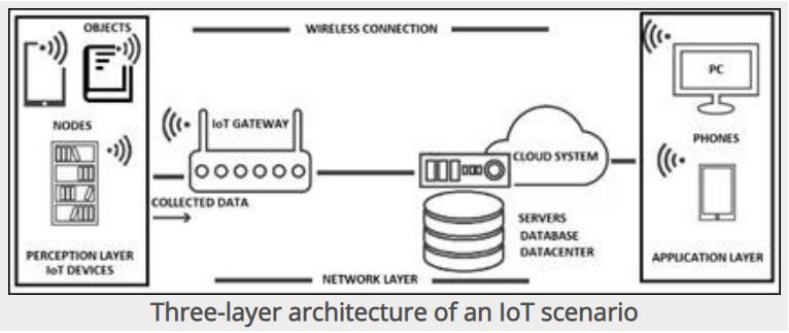


Figure 4. *Three-layer architecture of an IoT scenario.*

A a simple architecture of an IoT scenario which can be divided in three layers: Perception layer, Network layer and Application layer. Based on three-layer IoT architecture, the events are sensed by the nodes and after a local processing of the data; information will be sent to a cloud system server through the gateway (base station). The perception layer is the physical layer; it includes the sensors (RFID, WSN etc.), the event parameters or identity existence of the object. The network layer provides connectivity between the objects, network devices, wireless or cable connections, cloud system, as well as transmitting and processing the locally obtained data. It also includes the gateway component to receive the data sensed from the perception layer. The application layer is in charge of providing applications and services to the human or non-human users (i.e. Machine to machine case). It can specify various processes, programs, and applications in which the IoT can be positioned as a smart library management system.

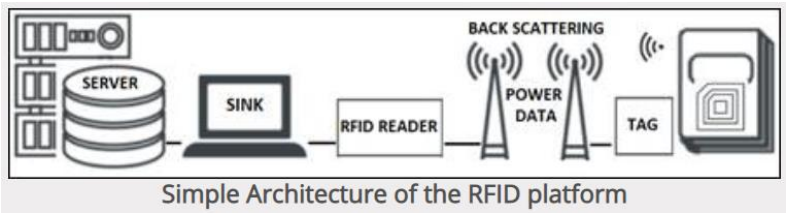


Figure 5. *Simple architecture of the RFID platform.*

A simple architecture of the RFID platform. A hand-held reader receives the signal from the RFID tag processing unit and then sends the data related to the signal to a cloud system. The WSNs refer to a group of small electronic sensors (green points), that can deploy over a region (sensor field) to sense, detect the event (red point) and gather the sensory data in the base station (or sink) in order to monitor the area and make a decision related to that event. The sensor nodes in WSNs can build a non-centralized network architecture

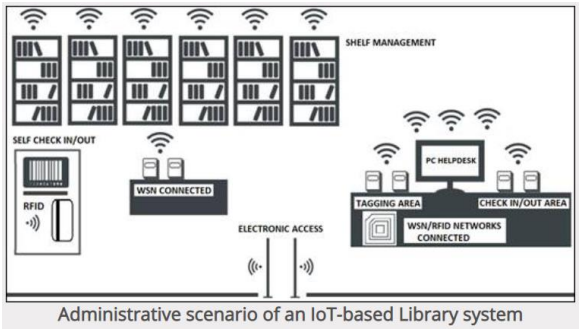


Figure 6. *Administrative scenario of an IoT-based Library system.*

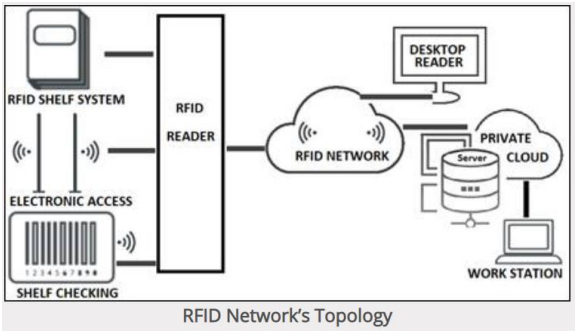


Figure 7. *RFID Network's Topology.*

The main IoT library system scenario consists of some areas with their administrative divisions. Technologically, the proposed scenario for IoT library system is a combination of RFID and WSN networks. The first part of the scenario is based on a RFID topology. The second part is based on the WSN network [11].

The main administrative zone is the tagging area. The new books will be delivered to the tagging area for registering and receiving the RFID labels. The users hold the RFID ID-card and the registered book takes a unique pasted tag. In order to identify the book-shelves a unique special tag is attached on them. When the book holders pass the electronic access gate control, a reader receives the signal of a tag passed on the book as well as the ID-card.

The second part of the IoT Library system is the WSN-based distribution. A small mote (WSN sensor) is pasted together with the RFID label on each book. A mote is a wireless node in a sensor network which is capable of sensing, gathering and local processing of the information. These sensors (motes) will not be activated in case the books are located and locked in the shelves or if they are taken out of the library. They will be active just around the library place [7].

A wireless connection of the nodes is built when the RFID elements are activated and scanned by RFID reader. This can occur when the books (RFID tagged) are taken to the reading area or out of the library. The monitoring system can detect geographically those elements that are not in their physical places. This can avoid misplacing or loosing the mote-labeled element such as document or books. The WSN monitor system is enabled to physically detect the tracked objects and produces a detailed report log related to the geographical area or location of the objects. This system can be applied to books, other types of documents and the administrative staff to locate people geographically and in real time mode.

Thirdly, service for smart library:

Smart Services can be described as the application of the "spirit of innovation" of smart cities to the development of modern library services. Papers on smart libraries often focus on this solution, and present technological innovation as smart services, such as RFID, mobile

and wireless access, remote assistance, semantic web, artificial intelligence, the Internet of Things, machine translation, voice and image recognition, natural language processing, augmented reality for delivering new experiences in enjoying cultural heritage etc. Some papers describe smart library services as library-based ICT platforms, for document search, information retrieval, collaborative collection building etc [6]. Another characteristic of smart library services is the interoperability and interconnection with other information services. A smart library is an information hub connected with other libraries and urban services in a larger informational ecosystem. However, these innovative tools and services are smart only insofar as they are user-friendly and user-centered. Smartness means that the development of new tools and services is based on the assessment of real usage. The user defines the library. "Smart is more user-friendly than intelligent" [13]. Instead of trying to adapt the user to existing library services, smart libraries are required to adapt themselves to the user needs. The assessment of real usage can include mobile crowd sensing in order to support smart mobility, usage of library space, and access to library facilities [18], agile management, UX design, and personalized information discovery based on recommendations (algorithms).

3. CONCLUSIONS AND RECOMMENDATIONS

Conclusions

Ideas about the Web 5.0 era have been discussed since 2013, when we were not yet in the world of Web 4.0. And now in the context that Industry 4.0 has existed, Web 4.0 and Library 4.0 must also develop accordingly. The development of the library is associated with the development of technology, Once we have discussed the "library generations", through the stages of Library 1.0, 2.0, 3.0, of course we must go to Library 4.0. [2, 4].

Our idea is that, in the urban environment and on the university campus, "smart library" may be one major candidate for this new name. Traditional marketing says, without customers, no service. Setting this rule on its feet, we would rather say that without smart services, no smart people. The smart library is an option to stay in the game and defend the fundamental library assets as a "civic landmark" in the upcoming smart society [12, 15].

Proposing smart library model and information resources sharing for smart cities and the above analysis may not be comprehensive but one thing is for sure 4th generation library is smart library, based on on the Internet of Thing platform, symbiotic websites, connecting people via 3D virtual world, with large data sources, showing human states through artificial intelligence, expressions, emotions on the face, changing through real time, as well as support from open source software, content technology, cloud computing services, presenting scientific works in the form of fine arts, visual arts photo ... is the library of the future.

Recommendations

Further studies and discussions on smart libraries:

- Library 4.0 in universities, agencies and cities
- Advanced technology in library 4.0

- Connect everything in the library
- AI in the library
- ...

The authors hope to continue to be reviewed by experts, scholars and scientists in the future.

REFERENCES

[1] Le Ba Lam, Nguyen Hong Minh, Research on the development of the Web together with "Library Generations" and propose a model of Vietnamese university libraries in the context of Industry 4.0.

[2] Nguyen Hoang Son, Le Ba Lam, Hoang Van Duong (2019), Smart Library Generations (1990-2025).

[3] Kwanya, T., Stilwell, C. and Underwood, P.G. (2013). Intelligent libraries and apomediators: Distinguishing between Library 3. 0 and Library 2.0 // Journal of Librarianship and Information Science. No. 45 (3). - P. 187-197.

[4] Web 1.0 vs Web 2.0 vs Web 3.0 vs Web 4.0 vs Web 5.0. Flat World Business - WordPress.com.

[5] Aghaei, S., Nematbakhsh, M. A. and Farsanim, H. K. (2012). Evolution of the world wide web: From web 1.0 to web 4.0 // International Journal of Web & Semantic Technology. No. 3(1). - P. 1-10.

[6] Joachim Schöpfel (2018). Smart Libraries: *Infrastructures;* GERiiCO Laboratory, University of Lille, 59653 Villeneuve d'Ascq, France.

[7] The Importance of the Internet of Things (IoT) for Project Management, https://www.inloox.com/company/blog/articles/the-importance-of-the-internet-of-things-iot-for-project-management/

[8] Younghee Noh. (2015). Imagining Library 4.0: Creating a Model for Future Libraries // The Journal of Academic Librarianship. - 2015. - No. 41. - P. 786-797.

[9] Zaigham Mahmood. (2018). Smart Cities Development and Governance Frameworks.: Springer.

[10] Hubert C (2018), Smart Library and Smart Campus (Journal of Service Science and Management).

[11] Majid Bayani (2017), IoT-Based Library Automation and Monitoring system: Developing an Implementation framework of Implementation.

[12] Smart Libraries Newsletter (2018). Available online: https://journals.ala.org/index.php/sln/index.

[13] Nam, T.; Pardo, T.A.(2011). Conceptualizing smart city with dimensions of technology, people, and institutions. In Proceedings of the 12th Annual International Digital Government Research Conference: Digital Government Innovation in Challenging Times, College Park, MD, USA[CrossRef].

[14] Vanolo (2014). A. Smartmentality: The Smart City as Disciplinary Strategy. Urban Study

[CrossRef]

[15] Hauke, P.; Latimer, K.; Werner, K.U. (Eds.) (2013). The Green Library: The Challenge of Environment Sustainability. IFLA Publications 161; De Gruyter: München, Germany; ISBN 978-3-11-030972-0.

[16] Harrison, C.; Eckman, B.; Hamilton, R.; Hartswick, P.; Kalagnanam, J.; Paraszczak, J.; Williams, P. (2010). Foundations for Smarter Cities. IBM J. Res. Dev [CrossRef].

[17] Neirotti, P.; De Marco, A.; Cagliano, A.C.; Mangano, G.; Scorrano, F. (2014). Current trends in Smart City initiatives: Some stylised facts. Cities[CrossRef].

[18] Aithal, P.S. (2016). Smart Library Model for Future Generations. Int. J. Eng. Res. Mod. Educ. Available online: https://ssrn.com/abstract=2822978.

A SOLUTION OF MANUAL TOLL COLLECTION (MTC) & ELECTRONIC TOLL COLLECTION (ETC) SYSTEM FOR VIETNAM EXPRESSWAY

Mai Vinh Du[1*]

[1]*Campus in Ho Chi Minh City, University of Transport and Communications, No. 450-451 Le Van Viet Street, Tang Nhon Phu A Ward, District 9, Ho Chi Minh City, Vietnam.*

[*]*Corresponding author: duvm_ph@utc.edu.vn*

Abstract: The manual toll collection (MTC) and electronic toll collection (ETC) system have been applying for national road and expressway in many countries on the world from 1960s [5,6]. However, the MTC & ETC system have just applied from 2015s [1,2] for national road in Vietnam. This paper proposed a solution of closed MTC & closed ETC system for Vietnam Expressway which suitable with standards and regulations in Vietnam.

Keywords: manual toll collection (MTS) system, electronic toll collection (ETC) system, toll collection for expressway.

1. INTRODUCTION

In the Decision no.326/QD-TTg dated 1 march 2016 on approving the master plan on development of Vietnam's expressway network through 2020, with orientations toward 2030: The master plan sets up Vietnam's expressway network consisting of 21 routes with a total length of 6,411 km, including: 1) The North-South expressway: to consist of 2 routes with a total length of around 3,083 km, 2) The expressway system in the northern region: To consist of 14 radial routes linked with Hanoi capital with a total length of 1,368 km, 3) The expressway system in Central Vietnam and the Central Highlands: To consist of 3 routes with a total length of 264 km, 4) The expressway system in the southern region: To consist of 7 routes of a total length of 983 km, 5) The system of high-speed ring roads of Hanoi and Ho Chi Minh City is 713km long.

In parallel with the construction of the expressway network, the toll collection system will be designed and installed for returning on investment [6]. In initially time in Vietnam, some technologies applied for toll collection system such as DSRC 5.8GHz active technology (Ho Chi Minh – Dau Giay Expressway). But, in early 2016 the Ministry of Transportation has agreed to use the RFID passive technology 860 MHz-960 MHz (ISO/IEC18000-6C) for all MTC and ETC system on Vietnam national road and Expressway [3].

However, there are not enough detailed design standards & regulations on MTC and ETC system for Vietnam expressways now. From this fact, this paper proposed a solution of closed MTC & ETC system for Vietnam Expressway.

2. SELECTION OF TECHNOLOGY

2.1. Technologies in MTC system

Currently, there are some technologies used in MTC systems in the world and in Vietnam such as barcode ticket technology [1], magnetic stripe card technology, and smart card technology (IC-card) [2].

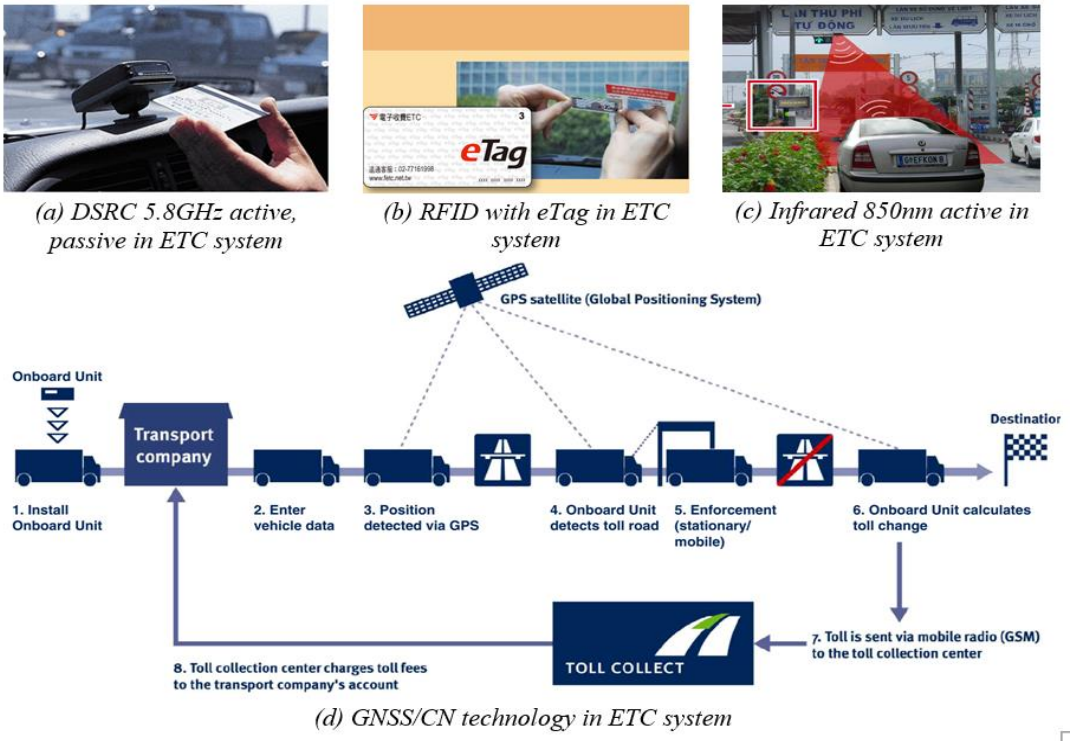
In the above cards, IC-card has many advantages (it is very easy to write, delete or rewrite the new information/data into IC-card [2]), so the IC-card can be reused in the closed MTC system on expressway. Therefore, we used IC-card for closed MTC system on Vietnam Expressway [2].

(a) Thẻ mã vạch (barcode) *(b) Thẻ từ (magnetic)* *(c) Thẻ thông minh IC-card*

Figure 1. *Some technologies in manual toll collection (MTC) system.*

2.2. Technologies in ETC system

Some technologies are used such as DSRC 5.8GHz passive & active, DSRC 5.9GHz, DSRC 2.45 GHz, RFID passive 860 MHz-960 MHz (ISO/IEC 18000-6C), Infrared 850nm active), Global Navigation Satellite System (GNSS) combined with mobile network (CN) is called GNSS/CN technology [4,5].



(a) DSRC 5.8GHz active, passive in ETC system *(b) RFID with eTag in ETC system* *(c) Infrared 850nm active in ETC system*

(d) GNSS/CN technology in ETC system

Figure 2. *Some technologies in electronic toll collection (ETC) system [4,9].*

RFID passive 860 MHz-960 MHz was experimented in 2015 in Vietnam. In early 2016 the Ministry of Transport has agreed to use an RFID passive 860 MHz-960 MHz for all MTC and ETC system on Vietnam national road and Expressway. The advantages of this technology are low cost, open standards, easy in application, there are many suppliers. On July 21, 2016, the Ministry of Transport issued Decision No. 2255/QD-BGTVT on common technical requirements for electronic toll collection system using RFID technology in road transport. Therefore, RFID passive 860 MHz-960 MHz is used in this paper [3].

2.3. Vehicle identification & classification technology (both MTC and ETC)

Automatic License Plate Recognition (ALPR) technology at entrance & exit lanes: ALPR which suitable for both day and night, two cameras used for MTC+ETC lane: 1 camera installed on gantry to capture number plate when ETC vehicle enters the lane, and 1 camera installed on toll island to capture number plate of MTC vehicle enters the lane. The camera on the gantry take number plate of ETC vehicle and send this data (number plate) to server in center office and used for opening of barrier for a valid ETC vehicle if the E-tag fails [2].

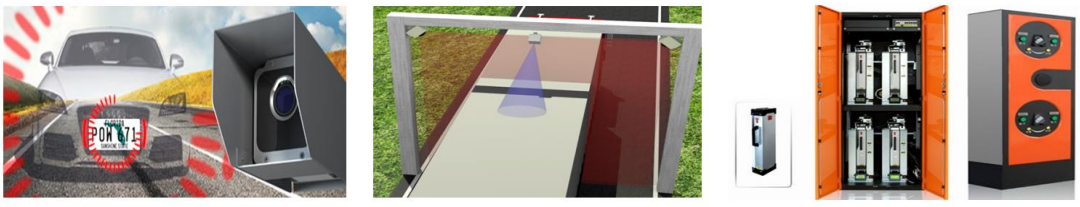
Vehicle identification/classification using laser technology at entrance & exit lanes: This technology has been applying in European, USA, Japan, ... In this proposed solution, we use a double laser scanner detector; each emitted laser beam makes a scan on 4 parallel planes with an angle of 96°. For each plane the sensor detects 240 points and it is able to detect the vehicle's profile with high accuracy. It is able to distinguish more than 20 classes of vehicles including motorbikes, cars, trucks, buses, containers ... (European standards) [2].

2.4. Automatic IC-card dispenser at MTC entrance lanes

To issue IC-card for closed MTC system at entrance lanes with high-speed, we use an automatic IC-card dispenser at entrance lanes with flowing functions: big IC-card tray capacity, it can be combined with camera, barrier, loop detector, ... The central processor is able to connect to the toll data server, ETC system, classification system to control equipment at lane. IC-card dispenser is also able to combined to ALPR camera and number plate will be pasted in the IC-card [2].

2.5. POS technology for fee collection at MTC exit lanes

To improving speed, accurate, convenient of fee calculation, we use an advanced POS machine at the exit lanes for cash payment. In addition, POS machines also can be integrated bill printers, IC card readers, non-cash payments such as ATM, Debit/Credit, electronic wallets (QR Code), NFC, mPOS [2], ...



(a) ALPR technology at entrance/exit lanes

(b) Laser technology at entrance/exit lanes

(c) POS technology at exit lanes

Figure 3. *Technologies in MTC and ETC system.*

With application of above technologies for the MTC and ETC system will have more convenient, more accurate for drivers and for fee calculation, and advanced POS machine at the exit lanes for cash payment also.

3. CLOSED MTC & ETC SYSTEM (FRONT-END)

3.1. Hardware architecture

Based on the selection of technologies in above, author proposes a system architecture for closed MTC and ETC system for Vietnam expressway as Figure 4 [2]. The proposed MTC and ETC system is flexible, upgradable, expandable in the future, able to integrated with other systems, highly compatible & not dependent on a unique supplier.

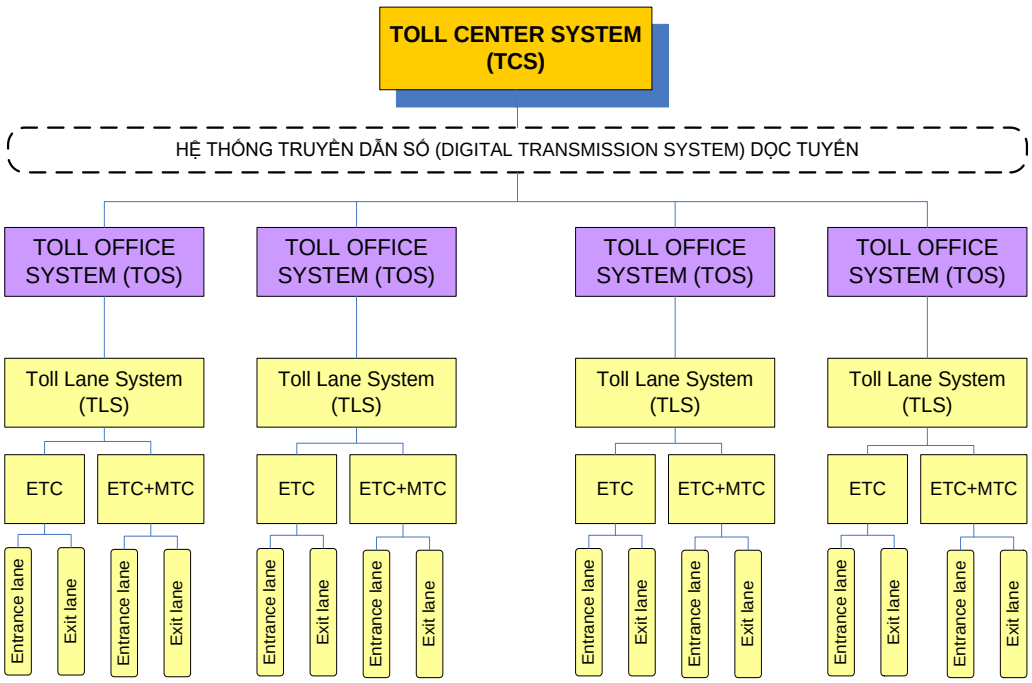


Figure 4. *Hardware architecture of proposed closed MTC and ETC system (Front-End).*

+ ***Toll Centre System (TCS):*** Includes server system & station computers, the server system provides services for employees/staffs, and the station computers (operation computers) have main HMI (Humane Machinery Interface) for control and manage of whole toll collection system.

+ ***Toll office system & toll lane system (TOS & TLS):*** Including the system at the building office of each toll stations, the system at toll lanes and alarm light system, …

+ ***Equipment for users/customers:*** Including E-Tag RFID tags and related issues for users/customers (E-Tag RFID), all kind of IC-card and related IC-card management.

3.2. Software architecture

+ ***Toll centre software (TCS):*** Toll Centre System Software manage system parameters such as vehicle type, payment type, toll fee and deliver these parameters to Toll Office System (TOS). Toll Centre System receive toll collection transactions and customer data from TOS. The TCS have two

main components: 1) Central management software (TMS) manage all toll collection transactions and customer data from TOS and check all transactions at all toll stations along the expressway, 2) Operation monitoring software (OMS) has the main function of monitoring and reporting on toll collection and equipment activities on the whole expressway [2].

+ ***Toll office system software (TOS software):*** The Toll Office System Software provides functions to manage, support and supervise all toll collection transactions at each TLS. TOS software will receive system parameters and system categories from TCS. TOS also have function to manage all accounting activities at Toll Office. TOS also have function to monitor, to post-check all toll transactions received from Toll Lane System [2].

+ ***Toll lane system (TLS software):*** The TLS software provides functions to check and collect toll fee for vehicle travel on expressway. Main function of TLS software is to process toll transactions (entry, exit transactions) and send these toll transactions to TOS for cashing up, post-checking, etc. And support the implementation of abnormal transactions [2].

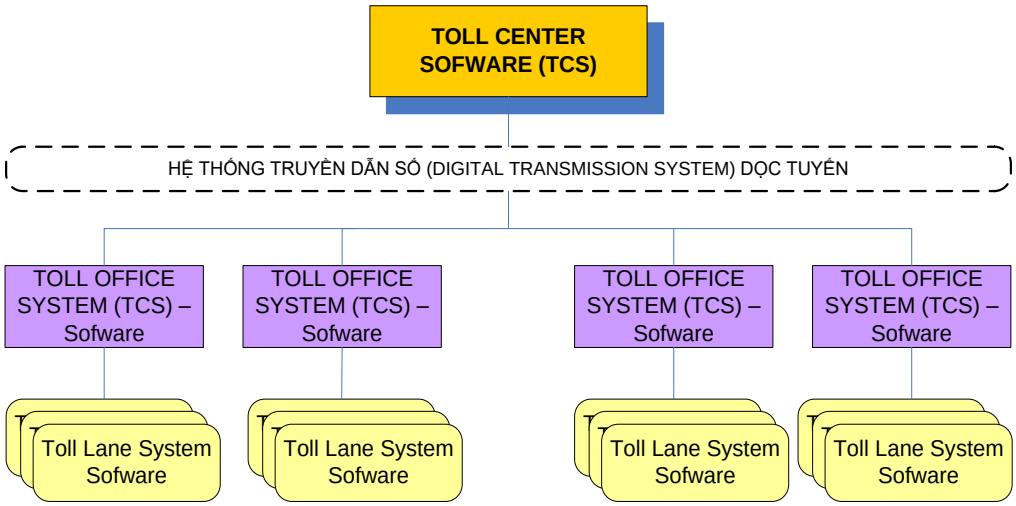


Figure 5. *Software architecture of MTC & ETC software (Front-End).*

3.3. Procedures of closed MTC system

In closed MTC system, vehicles entering the expressway and will stop at the entrance lanes to pick up the card from IC-card dispenser. And then the vehicles go to exit lanes and stop at the toll booth (POS machinery) for cash payment [2].

The procedures for closed MTC system at entrance lanes are as follows:

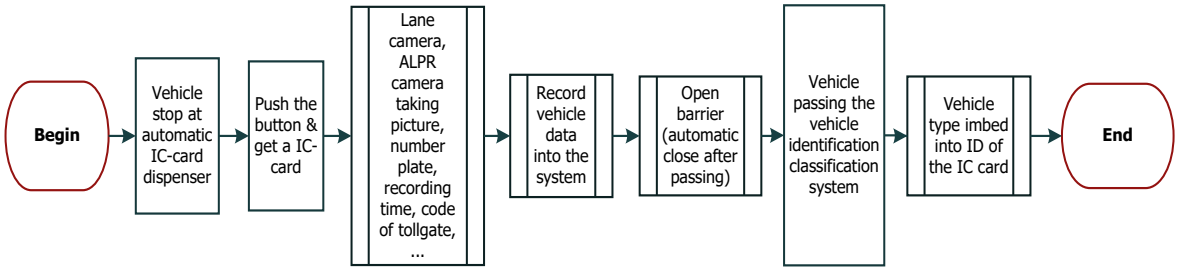


Figure 6. *The procedures for closed MTC system at entrance lanes.*

The procedures for closed MTC system at exit lanes are as follows:

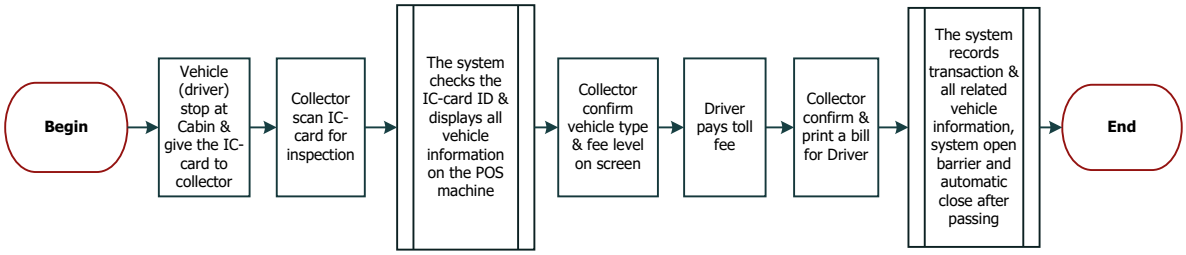


Figure 7. *The procedures for closed MTC system at exit lanes.*

3.4. Procedures of closed ETC system

Firstly, customers will be registered ETC system service at the toll office, E-tag dealer or at the registration center. Customers will be provided an RFID E-Tag including an ID account number created on the central system, this account information including (ID account umber, customer information, vehicle information, ID number of E-Tag (vehicle type, vehicle characteristics, license plate ...), time of issue, account information (current balance, payment history, history service, ...), validity period ... Registered RFID E-Tag will be pasted to the vehicle glass (if the glass contains metal, we past to the lamp) .E-Tag card provides information of ID account number to help the system identify customer accounts, while all information about customer accounts is stored and automatically updated to the Back-End [2].

Procedures of closed ETC system at entrance lanes are as bellow:

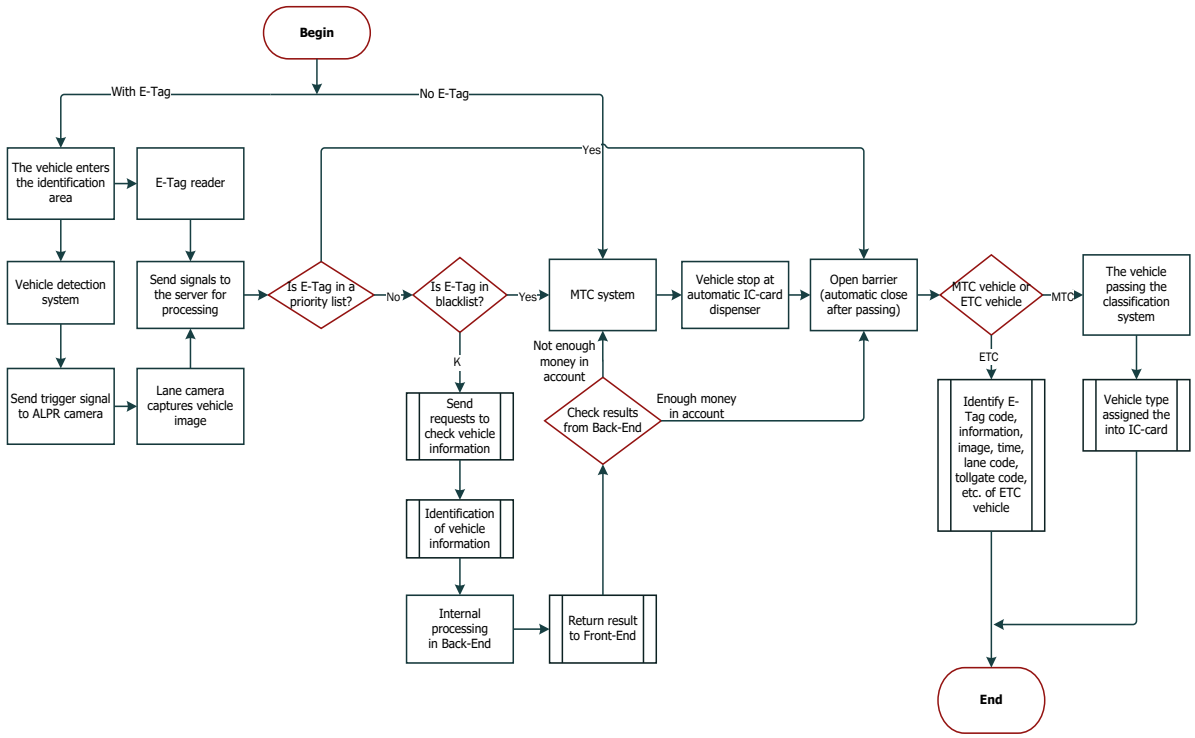


Figure 8. *Procedures of closed ETC system at entrance lanes.*

When the vehicles come to the exit lanes, the antennas & card readers at toll gate will

read the ID number of the E-Tag card via RFID communication. This ID number will be sent to the Back-End to verify account information: If the customer account is valid (the balance is sufficient for payment or is still valid), the ETC system will open barrier for vehicles, and notify Back-End of account deduction. If the customer account is not valid, the system will alarm and change to closed MTC procedures. At the same time, all information about the transaction will be sent immediately to the Back-End to update with the specific transaction number. Reconciliation, clearing, settlement between related units are carried out and updated into the database of the Back-End system at the station and card centre based on ETC transaction reports at the station [2].

Procedures of closed ETC system at exit lanes are as bellow:

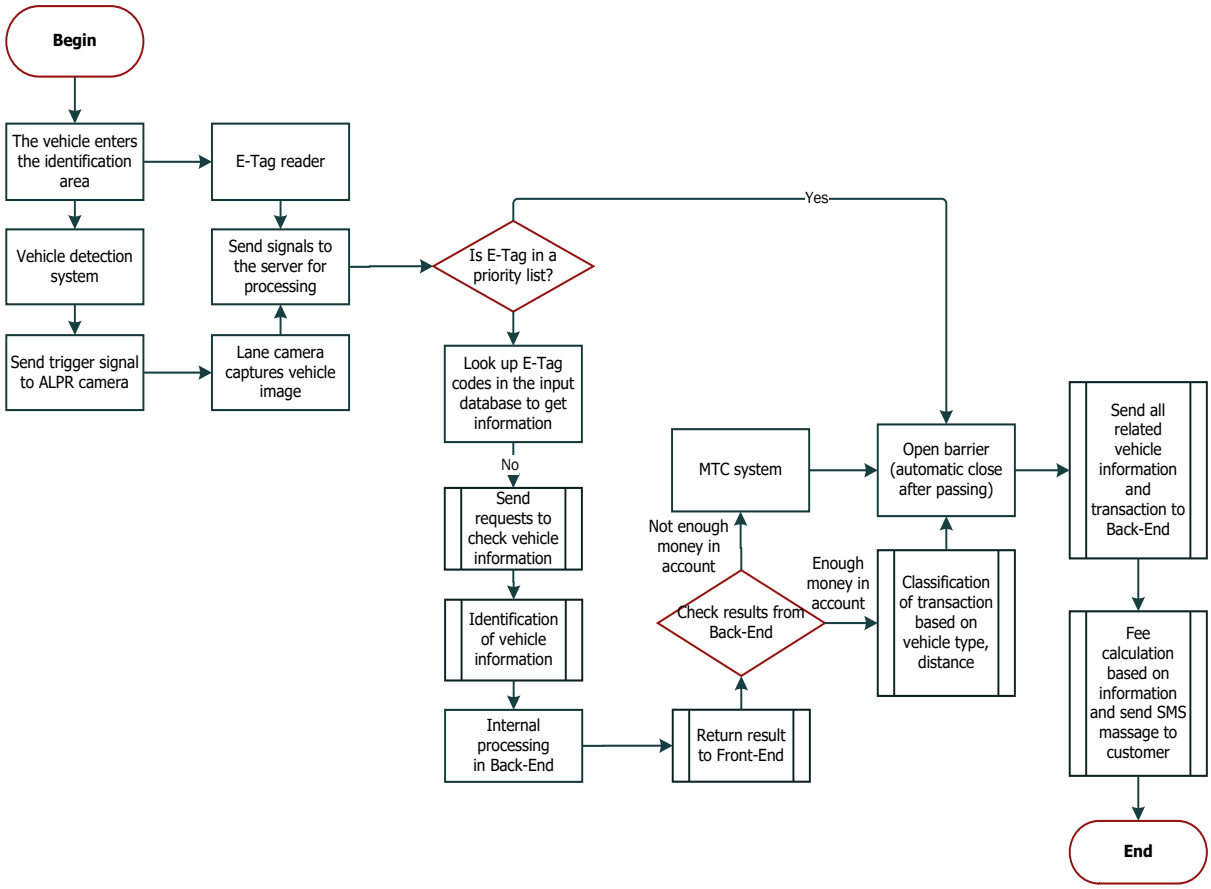


Figure 9. *Procedures of closed ETC system at exit lanes.*

3.5. Types of cards in closed MTC & ETC system

The employee ID card (IC-card): The employee ID card is used to identify all tollgates by connecting via IC card reader/writer devices at the beginning of the employee's work in the entire toll collection system on expressways, they are coded with an identification number and printed with an employee portrait [2].

Toll card for closed MTC system (IC-card): Customer will receive this IC-card from IC- card dispenser at the entrance lanes. Vehicles entering the expressway and will stop at the entrance lanes to pick up the card from IC-card dispenser. And then the vehicles go to exit lanes and stop at the toll booth (POS machinery) for cash payment [2].

Card used for ETC system (E-Tag): E-tag is a passive RFID tag provided by the ETC management Unit, customers will be provided an RFID E-Tag including an ID account number [2].

4. CONNECTION BETWEEN FRONT-END & BACK-END

As mentioned above, data exchange model between the Front-End (at toll plaza) and the Back-End (ETC service provider) will be complied with the Decision No. 2255/QD-BGTVT about the announcement of general specification on ETC (Electronic Toll Collection) system using RFID technology in road traffic and transport on July 21, 2016 of the Ministry of Transport [2,3].

4.1. Data exchange model

Data exchange model between the Front-End (at toll plaza, entrance station B and exit station B) and the Back-End (ETC service provider) for closed ETC system as shown in the following figure [2,3]:

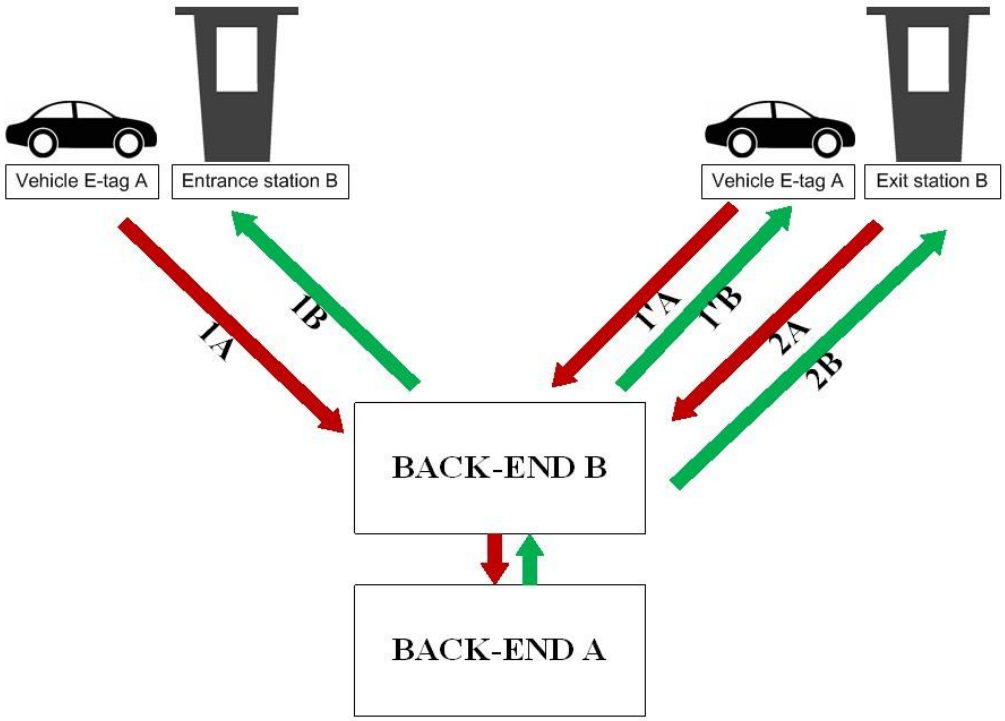


Figure 10. *Data exchange model between the Front-End (at toll plaza, entrance station B and exit station B) and the Back-End.*

For more detail of data exchange description, specification and detail requirements between the entrance station B & exit station B (Front-End) & the Back-End for ETC system, please see more detail in [2,3].

4.2. Security mechanism & data transfer rate

The security mechanism and transfer rate must be considered for data exchange between the Front-End (at toll plaza, entrance station B and exit station B) and the Back-End

(ETC service provider) for closed ETC system as requested in Decision No. 2255/QD-BGTVT, we need a high speed transmission line for data exchange with back-up line for urgent case (data transmission problem), the time for sending & receiving a response between Front-End and Back-End is no more than 200 ms [2,3].

4.3. In case of disconnections

If there is a problem with the transmission line between TLS and TOS, the TLS software will automatically switch to offline mode. In this mode, TLS will convert the entire connection with the TOS (TOS data server, ALPR module ...) into an internal mode. Therefore, all toll transactions will be stored on the internal server in lane, images in offline mode will be sent to the TOS when the transmission line is restored, the supervisor will check the transactions based on ALPR images without license plate [2,3].

5. CONCLUSION

The paper proposed a solution of closed MTC & closed ETC system for Vietnam Expressway which suitable with standards and regulations in Vietnam. And the closed MTC & closed ETC system on the expressway give us more convenient, more safety, more effectively in expressway management and operation, … and reducing of travel time.

REFERENCES

[1] TCVN 10849:2015 do Ban soạn thảo Tiêu chuẩn kỹ thuật quốc gia về Hệ thống thu phí điện tử biên soạn, Bộ Giao thông vận tải đề nghị, Tổng cục Tiêu chuẩn Đo lường Chất lượng thẩm định, Bộ Khoa học và Công nghệ công bố, 2015

[2] Hồ sơ thiết kế thi công hệ thống thu phí kín (ETC, ETC+MTC) cho cao tốc Bắc Giang – Lạng Sơn, Công ty TNHH tư vấn thiết kế công nghệ thông minh lập, 2018.

[3] Quyết định số 2255/QĐ-BGTVT về việc ban hành yêu cầu kỹ thuật chung đối với hệ thống thu phí điện tử tự động không dừng sử dụng công nghệ RFID trong giao thông vận tải đường bộ, ban hành ngày 21/7/2016.

[4] J.G. Jordán, F.Soriano, A comparison of different technologies for EFC and other ITS applications, in Proceedings of Intelligent Transportation Systems, 2001, pp. 1172-1177.

[5] A. Ashok, 2 V. Thangavelu, 3N. Harikrishnan, Electronic Toll Collection using Active RFID System, in Georgian Electronic Scientific Journal: Computer Science and Telecommunications, 2010, No.1(24), pp.5-12.

[6] G.K. Andurkar, V.R. Ramteke, Smart Highway Electronic Toll Collection System, in International Journal of Innovative Research in Computer and Communication Engineering, Vol. 3, Issue 5, May 2015, pp.4087-4094.

BUILDING A REAL-TIME CARS FLOW MONITORING SUBSYSTEM USING RASPBERRY

Trinh Luong Mien[1*], Vo Van An[2]

[1]*University of Transport and Communications, No.3 Cau Giay Street, Lang Thuong Ward, Dong Da District, Hanoi, Vietnam.*

[2]*Binh Duong Economics and Technology University, No.530 Binh Duong avenue, Thu Dau Mot, Binh Duong, Vietnam.*

[*]*Corresponding author: mientl@utc.edu.vn*

Abstract: The cars flow monitoring on roads or at intersections is an important traffic problem, applied in many subsystems of the intelligent traffic system - the advanced transport architecture in the era of industry 4.0. This article offers a new solution for integrating hardware and developing software of the cars flow monitoring subsystem using raspberry and camera. The raspberry acts as a data processing station, is mounted on the roadside pole, installed with the images data processing software, and communicated wirelessly to the supervision centre, and then to provide to the managers, drivers with information about traffic, location of cars on the route, via website or mobile apps. The experimental results on the roads show that the proposed cars flow monitoring subsystem meets the technical requirements, has high economic efficiency, brings the added value to users.

Keywords: cars flow monitoring, vehicle, intelligent traffic system, raspberry

1. INTRODUCTION

The cars flow monitoring is an important traffic problem in the road intelligent traffic system. The vehicle flow is an important data to support for many problems, such as control, supervision and inventory of vehicles on the road. At intersections with the signal lights, vehicle flow is an important input parameter to appropriately control signal time to suit the density of vehicles on each lane [1,7,8,18], which contributes to reducing traffic congestion. At road toll stations, vehicle flow is an important data to regulate traffic lanes to avoid traffic congestion [1,13-18], as well as to support the inventory management [1,18]. Vehicle flow on the streets/roads is an important basis for the government management unit to make effective policies on the management and development of road transport infrastructure [3,8].

In the era of industry 4.0, the technology to monitor the cars flow on the roads has developed strongly. Today, the cars flow monitoring subsystem not only provides managers with the typical car and quality of cars on the road in real time, but also provides the cars' status and location and signs ... on the road. This module also helps managers proactively support and timely rescue vehicles participating in traffic when incidents occur on the road. At the same time, this module also provides the driver with information about the density of vehicles traveling on each road, from which there is a safe and effective driving plan.

Studies on the vehicle flow monitoring have been published in many scientific works in

Vietnam [1-8] and in the world [9-18]. These studies have initially yielded positive results. However, some research results use old technology [2,3], some solutions are quite expensive [11-18], or complex [11,12], or designed for the specific applications [4-7,13-4], has not really brought benefits to users - providing real-time traffic information on the road remotely [4-6, 11-18].

In recent times, studies of vehicle flow monitoring have focused on the direction of using specialized cameras to take pictures of the current state of vehicles on the road, then transfer image data, video to the data processing centre, and from which give recommendations and warnings to the managers and drivers [4-6,8,11-18]. This solution requires large project investment because it requires a specialized computer with a strong configuration, high-speed data transmission, a camera with the sharp video and large image sizes.

Recently, raspberry is known as an embedded computer that allows good processing of image data and is becoming more and more popular [9,10]. Raspberry is about the size of a hand-phone, running an open Linux operating system, equipped with a powerful Broadcom BCM283x processor (integrated CPU, GPU, audio/video processor, I/O port integrated on chip), low power consumption, low cost (35USD) [9,10]. Raspberry allows configuration as a computer station that processes data collected from cameras in the field without transmitting data to the supervision centre (to process the image data)

Application of the advanced technologies in the era of industry 4.0, in order to bring the added value to users with low investment costs, a solution to build the cars flow monitoring subsystem in real time on the highways or at intersections using raspberry and camera, is proposed in this article. This research focuses on integrating hardware and developing software for the cars flow monitoring subsystem using raspberry and infrared camera. The raspberry is mounted on the roadside pole, installed the data processing software, communicated wirelessly to the supervision centre, and from which to provided to managers and drivers about the vehicle flow information on the route, via the website or mobile apps.

The organization of the paper consists of six sections. Section 1 presented research content. Section 2 proposed the structure of the cars flow monitoring subsystem. Section 3 deals with building hardware for the data processing station in the field using raspberry. Section 4 deals with developing software for the data processing station in the field, training network and install the proposed algorithm on the kit raspberry. Section 5 presents the experimental results. Finally, section 6 presents conclusions.

2. PROPOSED STRUCTURE OF THE CARS FLOW MONITORING SUBSYSTEM USING RASPBERRY

Raspberry-pi is a very small, low-cost, mobile-phone-size computer, running the Linux operating system, introduced in 2012 at Cambridge University. Raspberry-pi combined with the keyboard, monitor and 5VDC power can be operated as a normal computer, allowing use of office applications, listening to music, watching high-definition movies. Raspberry-pi uses SoC (system-on-chip), Broadcom BCM283x. The CPU, GPU, audio, communication components are all integrated into one chip system. This SoC-chip is located below the black 512MB Hynix memory chip in the middle of the board (Figure 1) [9,10]. Raspberry-pi is a computer equipped with an open Linux operating system that allows you to easily program and customize the user programming for different needs and purposes.

The research in this paper focused on building the images data processing field station (on the roadside pole) with high computing speed, small-size, low-cost, allowing communication of information data via GPRS network. The field processing station receives images from the field camera, then processes the data at the field, and sends the type/number of cars to the supervision center, and all information to the users/drivers via the website or mobile apps. However, this study only focused on cars traveling on highways or at intersections, cars with strong heat sources, like cars using internal combustion engines.

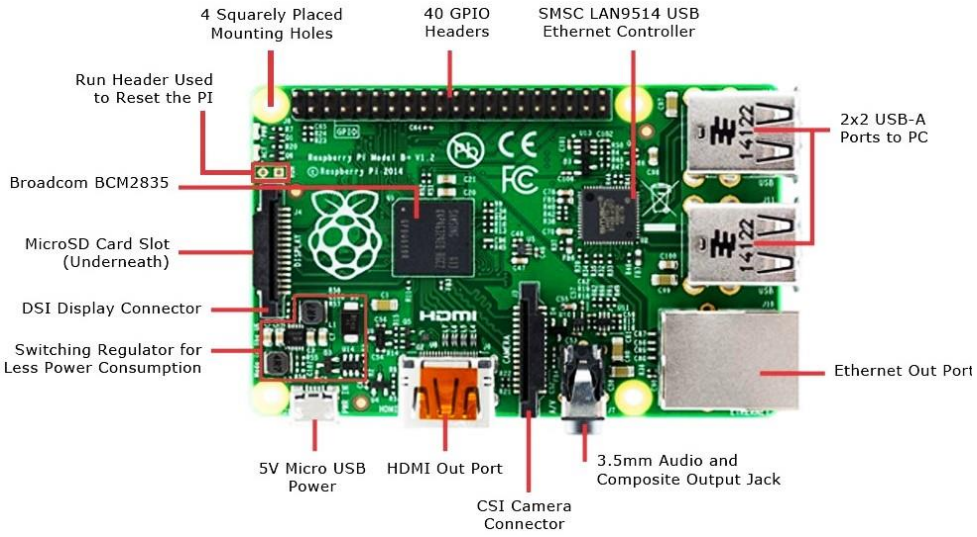


***Figure 1**. Kit raspberry-pi B v2.*

The structure block diagram of the cars flow monitoring subsystem in this study, is proposed as in Figure 2. The cars flow monitoring subsystem consists of three main parts.

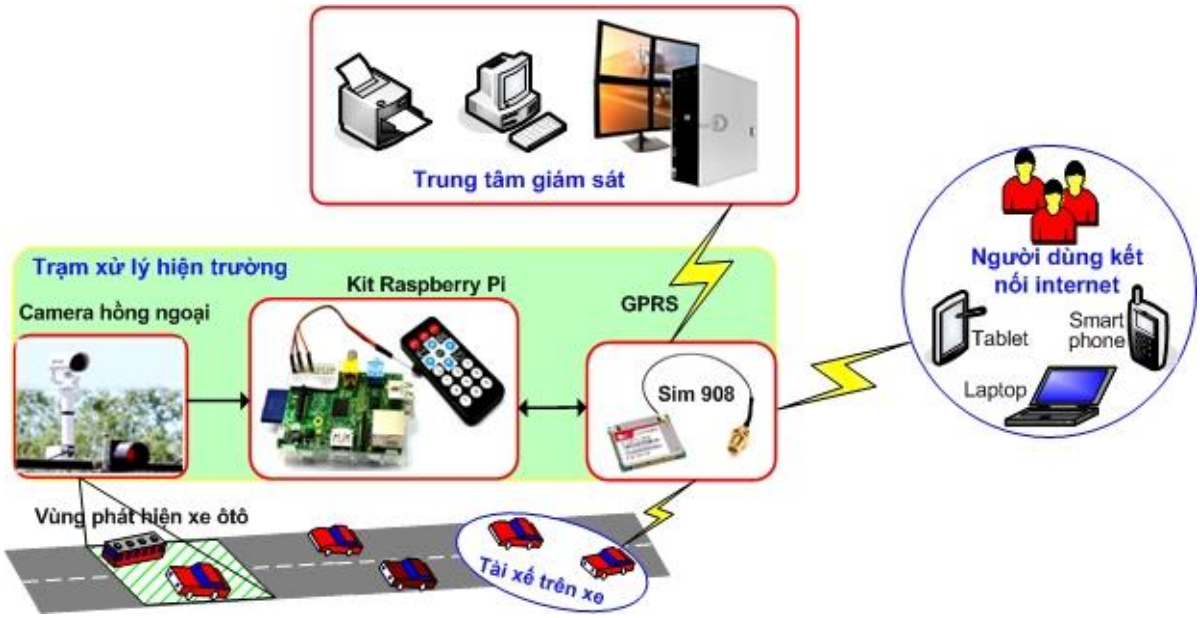


***Figure 2**. The block diagram of the cars flow monitoring subsystem.*

The first part is the field data processing station, including the kit raspberry-pi, integrated with SIM908 module supported GPS and GSM/GPRS network. Kit raspberry-pi is

connected to the day/night infrared camera. The main function of this station is to record video and take photo of the vehicles on the road, and then to process the image data received by the camera to classify vehicles and count the number of the vehicles on the road. Result of this processing is then transmitted the center, as well as notify the drivers/users by GPRS network. This is the main research content in this paper.

The second part is the supervision center which includes the server station, warning monitoring computers to support traffic conditions on the highways, and an intranet system with internet connection. The central station is equipped with traffic monitoring software on the route, allowing real-time traffic status information, online video observation of vehicles on the route, manage and update traffic status data automatically into the database, issue warnings to the managers, make decisions to coordinate traffic flows on routes, locate and decide to rescue vehicles with incidents on the route, make automatic reports according to time and event.

The third part is the warning block to drives/users. With electronic device connected to the internet, we can watch online video of vehicles from the field processing station. The users can receive warning information, watch video on smart phones or vehicle-specific monitors, via the website or mobile apps.

3. BUILDING HARDWARE FOR THE DATA PROCESSING STATION

The data processing station in the field is built based on the integrated kit raspberry-pi with SIM908 module, infrared camera and 5VDC power, key-pad, as shown in Figure 3.

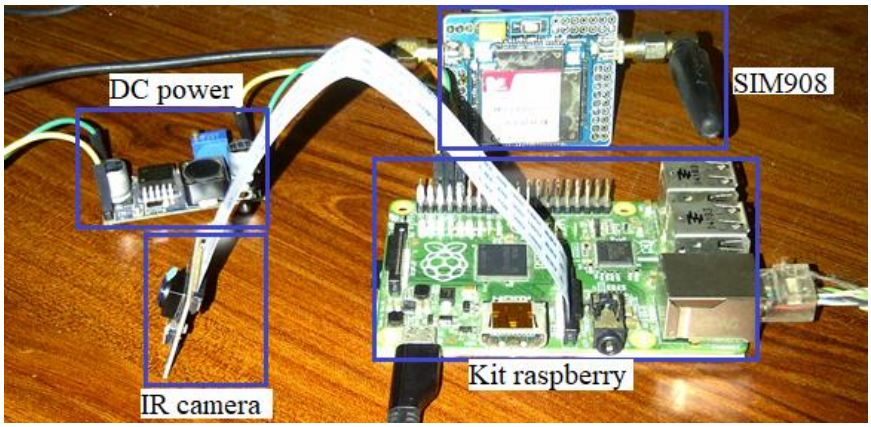


Figure 3*. The data processing station in the field using raspberry.*

The field station of this subsystem uses kit raspberry-pi. This raspberry kit is based on Broadcom BCM2835 processor, 700MHz, 256MB RAM, 32-bit RISC ARM processor cores, 5VDC power supply, very low power consumption of about 3.5W/h [9,10]. Raspberry is designed with lots of the I/O ports for connecting to various types of the I/O devices, such as Memory card slot; Micro USB power; TFT touch screen; Camera expansion; HDMI port; Ethernet port; Two USB ports; Stereo Audio port; TV port; GPIO port, SPI, I2C, control peripheral devices.

Module SIM908 receives satellite coordinate data which is sent to raspberry for processing. SIM908 connects to raspberry-pi by UART protocol, RX pin on SIM908

connects to TX pin on raspberry, TX pin on SIM908 connects to RX pin on raspberry, GND pin of SIM908 and raspberry connects to each other.

In this study, the authors use the infrared camera for testing subsystem. This camera is great for raspberry-pi, 5MP; 1.4µm x 1.4µm pixel with OmniBSI technology for high performance (high sensitivity, low cross-talk, low noise); Optical size of ¼ inches; Shoot or record videos with simple commands from the raspberry pi; Can control frame rate; 32bytes of embedded memory one time programmable.

The principle schematic circuit diagram of the DC power module is introduced in Fig.4.

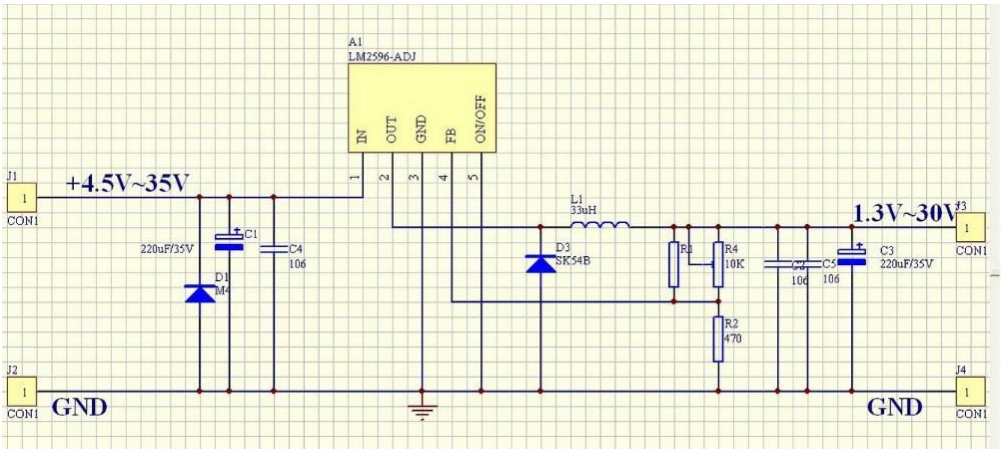


***Figure 4**. The principle circuit diagram of the DC power module.*

The principle schematic circuit diagram of the SIM908 module, communicated with raspberry, is presented in Fig.5.

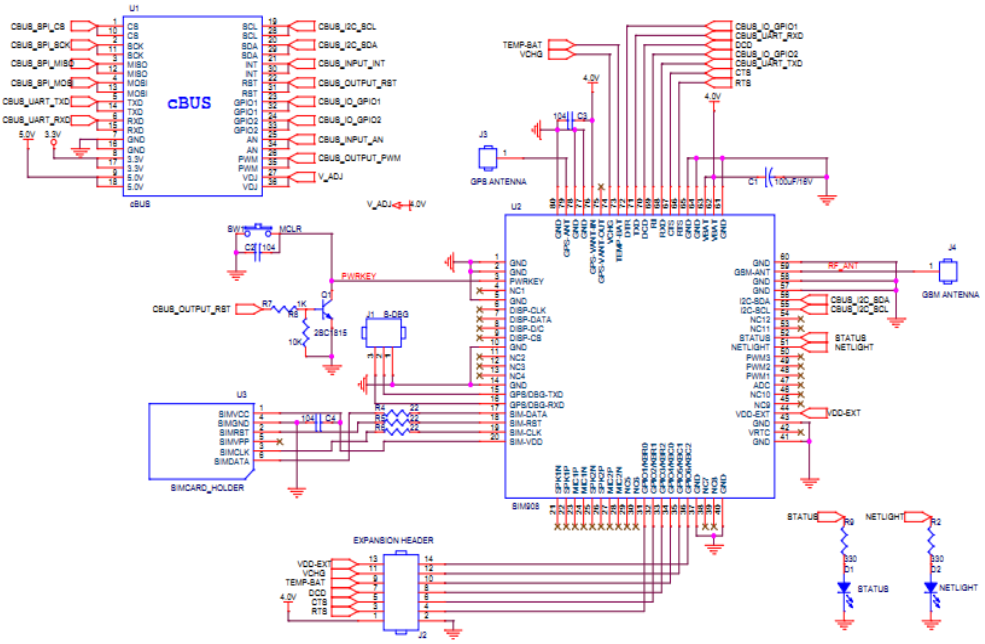


***Figure 5**. The principle circuit diagram of SIM908 module, communicated with raspberry.*

4. DEVELOPING SOFTWARE FOR THE DATA PROCESSING STATION

Raspberry supports many operating systems, such as Raspberry NOOBS, Linux Raspbian, RaspBMC for internet TV, RISC OS. However, Raspian operating system is popular and suitable for control and monitoring applications.

After downloading the Raspian operating system to the memory card and connecting to the kit raspberry-pi, the author executes the command Startx or Init5 to start the raspbian operating system, and then install the Python language for the kit raspberry-pi. After that, the authors installed the proposed algorithm to identify vehicles (cars, buses, trucks) passing through the surveillance area to mount the cameras on the roadside pole.

The proposed algorithm for the vehicle identification was developed based on the multi-task cascaded convolutional neural network (MTCNN) [18] with 3 stages of deep learning. Each stage uses convolution networks to analyze images, extract slope characteristics, detect edge and depth of each image area. Then compare each image area with the vehicle training data to find the area with the most data, which is the vehicle storage area.

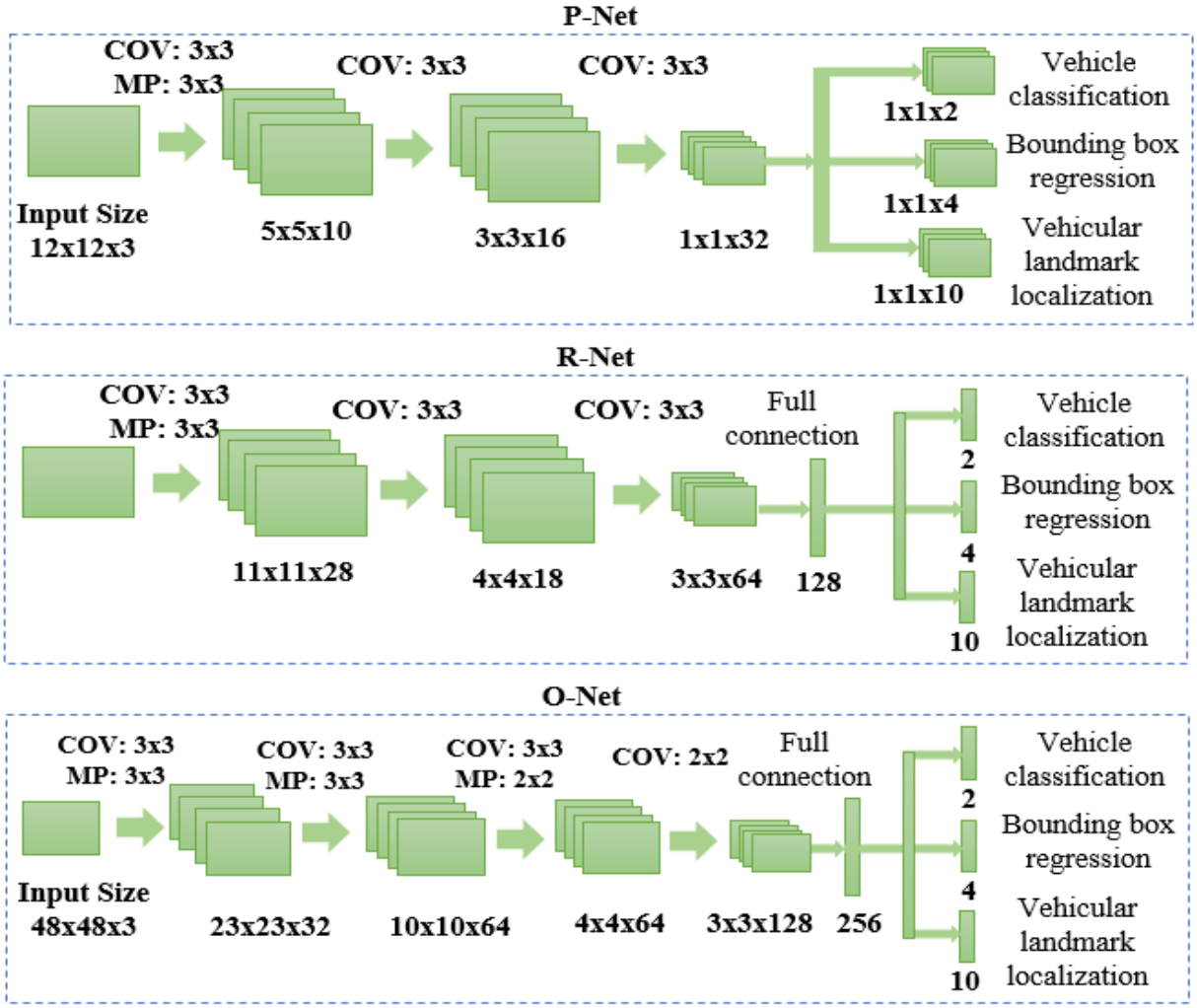


***Figure 6**. The proposed MTCNN-based vehicle feature detection method is divided into three stages: 1-proposal net, 2-refine net, and 3-output net.*

The three stages of the data processing in the proposed MTCNN-based algorithm as:

Stage 1: Putting pictures on the P-net network

+ Read images from videos to include in the program
+ Create multiple scaled copies of images
+ Insert scaled images into P-Net
+ Collect P-Net output
+ Clear bound boxes with low reliability
+ Convert 12x12 kernel coordinates to an image without scaling
+ Remove the kernels in each image rate
+ Converts bounding box coordinates to an image without scale
+ Reshape the limited box into a square

Stage 2: Putting data into R-net network

+ Insert scaled images into R-Net
+ Collect R-Net output
+ Clear bound boxes with low reliability
+ Maximum removal for all boxes
+ Converts bounding box coordinates to an image without scale
+ Reshape the limited box into a square

Stage 3: Bringing data to O-net and analyzing results

+ Insert scaled images into O-Net
+ Collect O-Net output
+ Clear bound boxes with low reliability
+ Convert boundary frames and landmark coordinates to an image without scaling
+ Maximum removal for all boxes

Provide the output results

+ Package all coordinates and confidence levels into one frame area
+ Returns the result as the coordinates of the car

The flowchart of the entire vehicle recognition process is given in Fig.7.

In order to install the proposed vehicle recognition algorithm on raspberry, the authors firstly need to train the proposed MTCNN network model in the article, and then to download the trained network model to the raspberry.

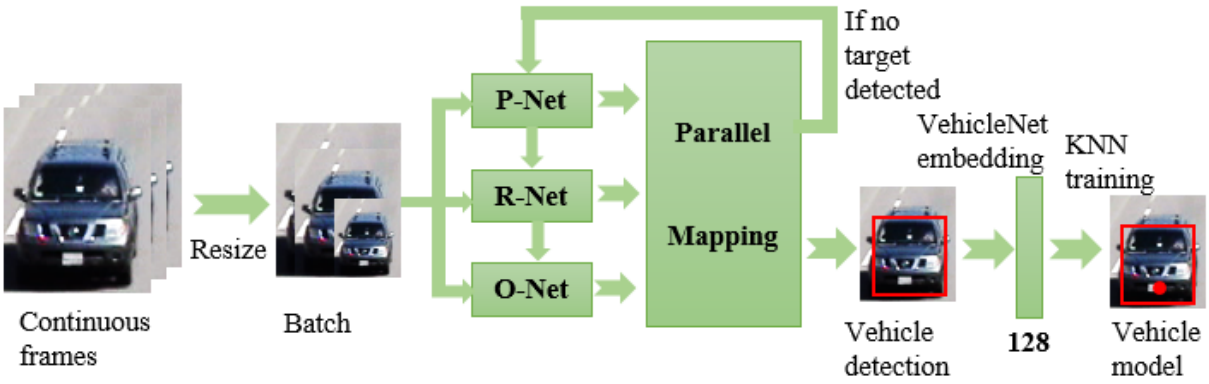


***Figure 7**. The flowchart of the entire vehicle recognition process.*

The transfer learning process for the network model requires to prepare the sample data set of cars (trucks, buses, cars). During network training, the weight of the network will be changed continuously according to the loss function.

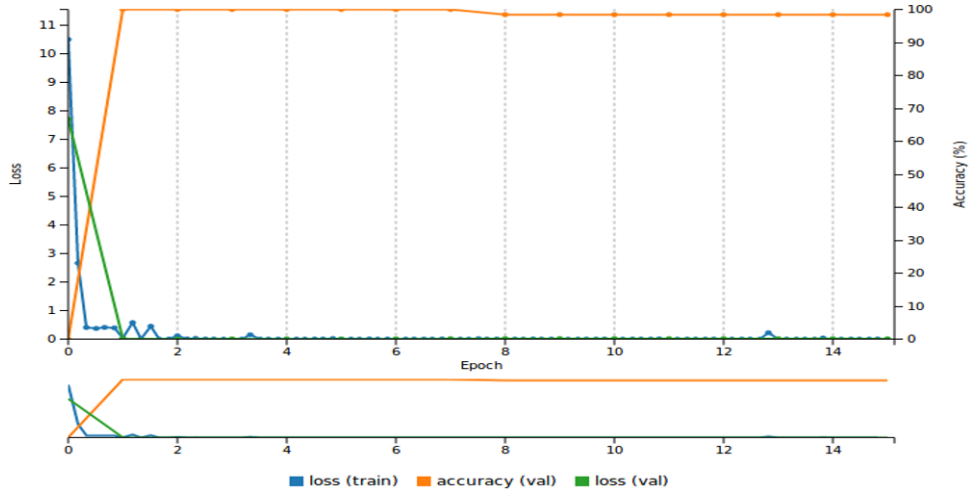


***Figure 8**. The result of training network process.*

The algorithm block diagram for recognition the vehicle, installed in raspberry, is presented in Figure 9.

The organization of the source code program on raspberry as below:

+ File cmakelists and folder build/: store the shell script to support for building source code on raspberry

+ Folder lib/: store Libraries MXNet and OpenCV to support for programming the vehicle recognition

+ Folder mtcnn_model/: store the models of the MTCNN algorithm for detecting the vehicle

+ Folder src/: store the main source code for the vehicle recognition

+ Folder image/: store the vehicle images for the vehicle recognition

+ Folder feature_model/: store the models insight vehicle - vehicular recognition

features

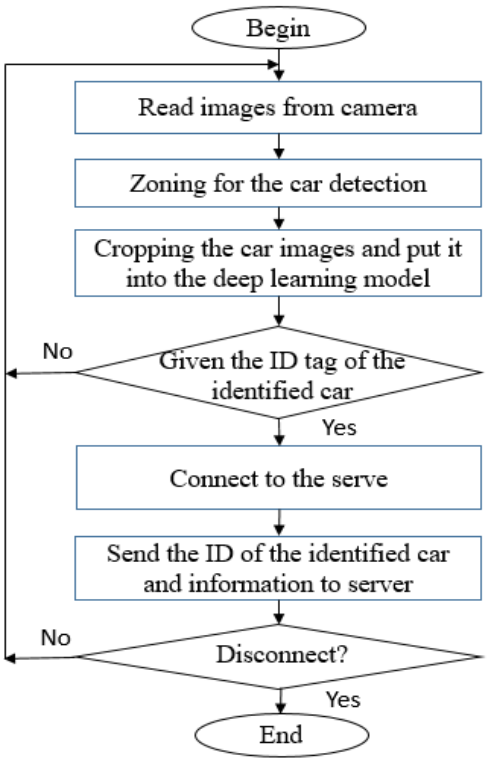


***Figure 9**. The algorithm block diagram for recognition the vehicle.*

5. SUBSYSTEM TESTING ON THE ROAD

The authors made the experiment of the cars flow monitoring subsystem on the Thang Long avenue, Hanoi, Vietnam, and obtained the results of the vehicle recognition, as shown Figure 10, Figure 11. The vehicle recognition model with high accuracy, about 90%, the images processing speed for a vehicle is about 0.5 second.

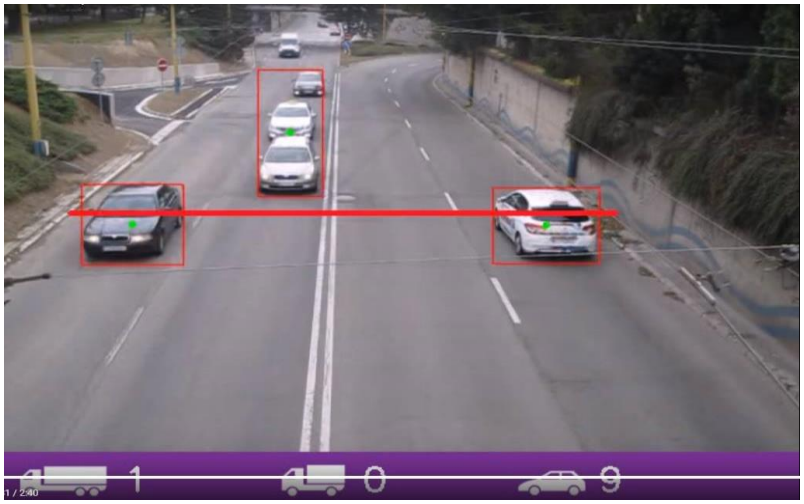


***Figure 10**. The result of the vehicles recognition at the field station on Thang Long avenue.*

The information on the number and type of the cars automatically sent to the supervision center from the data processing station at the field with the time cycle 5 seconds.

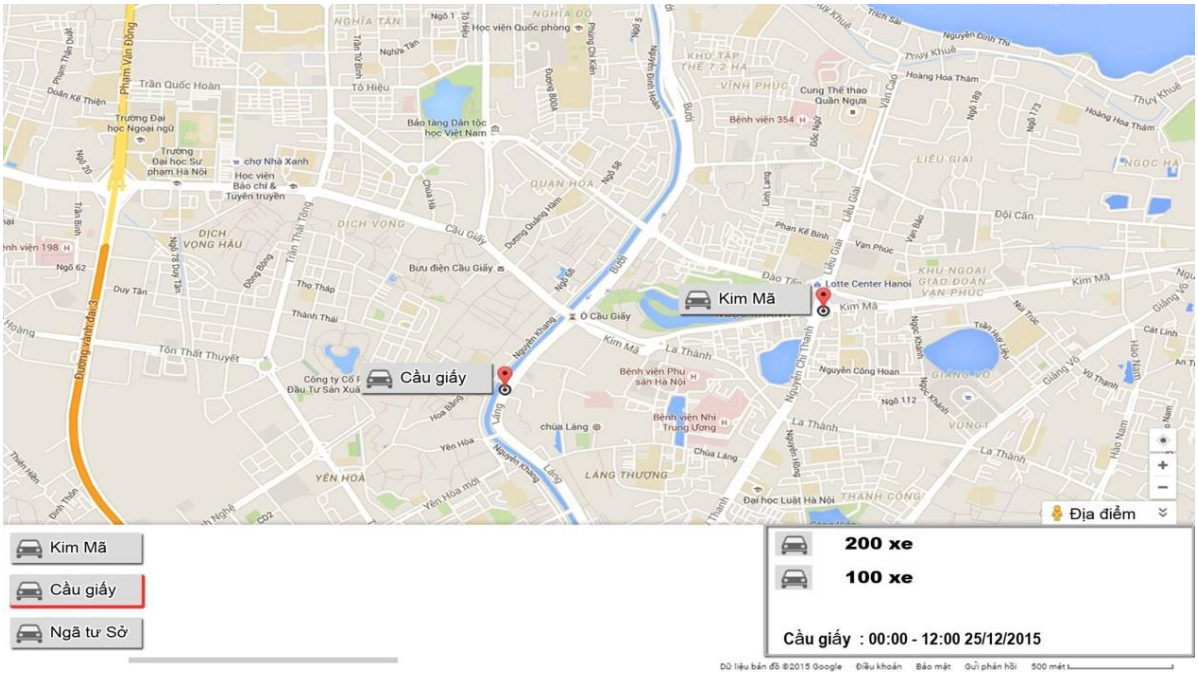


Figure 11. *The website based on Google map for supervising the vehicle on the streets.*

At the supervision center, through supervision software developed on Google map, we can know the number and type of the vehicles on the route, as well as their coordinates.

6. CONCLUSION

The article presents a simple, effective and low-cost solution to build the vehicle flow monitoring subsystem. This subsystem has the images data processing station located directly on the roadside pole, using raspberry and dedicated camera. Based on the multi-task model cascaded convolutional neural network, the authors have developed and installed the proposed recognition algorithm and counting the number of the vehicles on the kit raspberry. The authors have also built the website for supervising the vehicles flow at the center based on Google map.

The experimental results of the subsystem on the roads/streets (Thang Long avenue, Kim Ma str., Cau Giay str.) show that the proposed model for the data processing station in the field, is good and low-cost. Information on the traffic condition of the route is also easily accessed by mobile devices, by internet connection, through the website or mobile apps. The use of raspberry allows to provide timely and accurate field traffic information to users. This solution reduces the investment cost, stores simply the data at the center. This subsystem also provides information about the location of the incident on the map, thereby supporting timely and easy rescue.

Future research is to focus on perfecting vehicle recognition algorithms and designing and manufacturing the dedicated ARM-microprocessor boards for the data processing stations in the field, against the extreme weather and noise.

REFERENCES

[1] Le Hung Lan et al. Intelligent transportation system, volumes 1 and 1A, Transport Publisher, 2012.

[2] Le Hung Lan et al., Research and Application of Information Technology and Automation in Managing Urban Transport, National Science and Technology Development Research Project, KC-03-21, 2005.

[3] Le Hung Lan et al., Research, design, manufacture equipment, facilities and automated testing, monitoring, operating in service of road traffic safety, National Research on Science and Technology Development, KC.03.05/06-10, 2008.

[4] Pham Hong Quang, Building camera network with intelligent image processing system for traffic control and security monitoring, Software - Automation - Control Corporation, KC03.DA06/11-15, 2014.

[5] Le Quoc Anh, Phan Tuong Lai, et al., Application of real-time image processing technology in automatic traffic monitoring problem in Vietnam, 2005.

[6] Dang Quang Thach, Building an estimation algorithm of vehicle density on the road by image processing technology, Institute of Technology Application, 2016.

[7] Nguyen Chi Ngon. Research and design of intelligent traffic light systems, Journal of Science - Can Tho University, 2010.

[8] Dang Minh Tan, Nguyen Quoc Tuan and Do Van Duc. Image processing technology in organization, traffic management and traffic safety assessment, Proceedings of the Conference Development model of Smart Traffic Operation Center in Ho Chi Minh City, 2016.

[9] Jodi Jensen, Eben Upton, Raspberry Pi User Guide, Wiley Publication, 2012.

[10] Lady Ada, Introducing the Raspberry Pi 2 - Model B, Adafruit Industries, 2015.

[11] X. Chen, S. Xiang, C. Liu and C. Pan, Vehicle Detection in Satellite Images by Hybrid Deep Convolutional Neural Networks, Geoscience and Remote Sensing Letters, IEEE, vol. 11, no. 10, pp. 1797 - 1801, 2014.

[12] H. Cheng, C. Weng and Y. Chen, Vehicle detection in aerial surveillance using dynamic Bayesian networks, IEEE Transactions on Image Processing, vol. 21, no.4, pp. 2152 - 2159, 2012.

[13] Q. Tan, Q. Wei, S. Yang and J. Wang, Evaluation of urban road vehicle detection from high resolution remote sensing imagery using object-oriented method, Urban Remote Sensing Event, Joint, pp.1 C6, 2009.

[14] Saad Mohammed AI-Garni and Adel ben Abdennour, A Neural Network Based Traffic Flow Evaluation System for Highways, J King Saud Univ., Vol. 20, Eng. Sci., Riyadh, 2008.

[15] Mohamed Adel Al-Shaher, Detection and classification of vehicle types using machine learning technology, Periodicals of Engineering and Natural Sciences, Vol. 7, No. 3, September 2019, pp.985-815.

[16] M. N. Roecker et al., Automatic Vehicle Type Classification with Convolutional Neural

Networks, 2018 25th International Conference on Systems, Signals and Image Processing (IWSSIP), Maribor, 2018, pp.1-5.

[17] V. Vijayaraghavan, M. Laavanya, Vehicle Classification and Detection using Deep Learning, International Journal of Engineering and Advanced Technology, Volume-9 Issue-1S5, December, 2019.

[18] Nou Sotheany, Chaiwat Nuthong, Vehicle Classification using Neural Network, 2017 14th International Conference on Electrical Engineering, Computer, Telecommunications and Information Technology, 2017.

[19] Qilian Liang, Xin Liu, Zhenyu Na, Wei Wang, Jiasong Mu, Baoju Zhang, Communications, Signal Processing, and Systems, Proceedings of the 2018 CSPS Volume III: Systems, Lecture Notes in Electrical Engineering, Volume 517, Springer Nature Singapore Pte. Ltd. 2020.

ANALYSIS TRAFFIC CONFLICT TECHNIQUE FOR TRAFFIC SAFETY EVALUATION AT INTERSECTION

Nguyen Tuan Thanh[1], Cao Phuong Thao[1*], Vuong Xuan Can[1]

[1]*University of Transport and Communications, No 3 Cau Giay Street, Hanoi, Vietnam.*

[*]*Corresponding author: thaocp@utc.edu.vn*

Abstract: Traffic accidents occur frequently at intersections area because of conflict among vehicles. Many researches have been assessing traffic safety, however most of them are based on accident data that happened and caused injury as well as economic damage. Recently, conflict techniques provide a general view and solutions to prevent early collisions. This paper presents a method of conflict analysis using image processing technique and fuzzy comprehensive evaluation. The conflict parameters collected from two intersections Nguyen Chanh-Mac Thai To and Van Cao-Lieu Giai were processed and evaluated to give the solution of organization transportation at these intersections. The experimental results show that the proposed solution makes significantly decrease the dangerous conflict.

Keywords: traffic conflict technique, traffic safety evaluation, image processing

1. INTRODUCTION

Nowadays, the frequency and severity of crashes are two important data used commonly in assessing traffic safety on the road or in intersections. These datasets have two major disadvantages. First, the accident happened and caused the particular damage. Secondly, the process of collecting these data usually takes from 3 years to 5 years to achieve several samples that have sufficient reliability to analyze[1]. Assessing traffic safety based on traffic conflict techniques will be better for the shortcoming. Besides, the occurrence of traffic conflict is more frequently, so it will provide deeper and wider insight into failure that the last is an accident. Therefore, conflict parameters are considered as an alternative to collision parameters when study road traffic safety. Effective solutions could reduce human and material losses to the minimum by forecast and conflict situations analysis.

In the world, conflict techniques established to surrogate traditional road accident data were first introduced by Perkins and Harris[2]. According to published researches, the identified "traffic conflict" and the methods of how to recognize types of conflict were considered. Besides, parameters were set up to describe the process occurring in conflict scenarios and to modernization connection between conflicts and collisions[3,4].

In Viet Nam, there are few studies implemented to consider traffic conflicts, especially at signal intersections. Therefore, this paper exposes the application of conflict techniques and comprehensive fuzzy assessment methods to detect hazard levels of conflict points and its locations.

2. METHODOLOGY

2.1. FCE method

Fuzzy comprehensive evaluation method (FCE) also called fuzzy synthetic evaluation (FSE) is used to analyze and assess the system relied on fuzzy logic field and published by Zadeh [5]. The FCE method has also been widely used in many kinds of scientific domains involving traffic safety evaluation. Appling the FCE method is presented following:

Assume that domain U= (u1,u2,u3,...,um) denotes to set of assessment parameters. V = (v1,v2,v3,…,vp) is used to appoint a diversity of evaluation level or command control, so vk is a number that would be to express the link between evaluated factors and its consequence such as safety, less safety and dangerous. This Fuzzy relation is demonstrated in matrix R from U to V:

$$1. \quad R = (f_{jk})_{mxp} = \begin{bmatrix} f_{11}\ f_{12} \cdots f_{1p} \\ f_{21}\ f_{22} \cdots f_{2p} \\ \cdots\ \cdots\ \cdots\ \cdots \\ f_{m1}\ f_{m2} \cdots f_{mp} \end{bmatrix}$$

Where: denotes the membership degree j = (1,2,…,m) ; k = (1,2,…,p), where the appraisal object is measured vk as considering attribute uj; Rj = ( ) denotes a fuzzy set of attribute uj and $\sum f_{jk} = 1$.

Evaluation of traffic safety at a signal intersection based on conflict techniques and FCE methods would be implemented following 4 steps as below:

Step 1: Construction of domain of conflict indicators based on given safety level (I, II, III, IV)

Step 2: Construction of fuzzy membership functions of the indicators, based on its domain.

Step 3: Calculation of the weight of the indicators

Step 4: Determination of safety level based on the correlation coefficient

2.2. Data collection

Field data collection was conducted at 2 signal intersections in Ha Noi city in Viet Nam, including the Nguyen Chanh – Mac Thai To and the Van Cao – Lieu Giai – Doi Can intersections. The main reason for this selection is that the two intersections have many typical characteristics of Ha Noi's intersections. For example signalized four-leg intersections, mixed traffic, lack of some pedestrians and cyclists,…

Traffic data was collected at non-rush hour on weekdays by cameras. The recorded videotapes were used to play in the laboratory for achieving the information involving traffic flows and traffic conflicts. The camera scenes essentially focused on an area of inside intersection where frequently occur crossing conflicts. Besides, the research team also measured the intersection's geometric to calibrate the dimension in image processing

software. Data of crossing geometric features were collected due to the measured processing implemented by the Topcom GTS – 1002 total station.

The processing of video-data was implemented to achieve the value of different conflict parameters, include: time to conflict (TTC), conflict speed (CS) and deceleration rate (DR), which are common factors of traffic conflict techniques, are allocated as evaluation indicator[6]. The below section would demonstrate the main steps to analyzing the collected videos.

The video processing was performed following 3 steps as below:

Step 1: Initializing video in Tracker 8.0 software

Step 2: Setting calibration factors and coordinates

Step 3: Tracking the movement of the observed object.

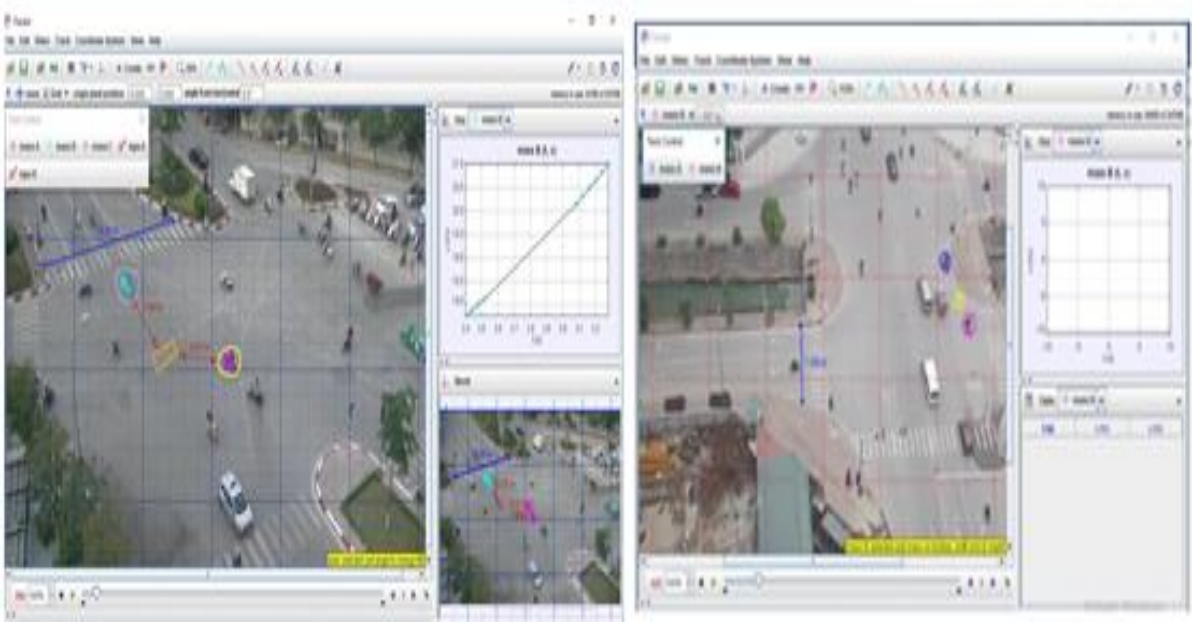


Figure 1. *Video processing by using Tracker software.*

3. RESULT

The number of traffic conflict points recorded at two selected intersections are 319 samples. At each point, there are three conflict factors collected by recognizing the conflict, predict area or position that two vehicles could collide if they remain at their speed and direction, and measure time to collision, conflict speed. These parameters were indicated by cumulative frequency curves, as shown in Figure below:

Data was recorded by video processing are input parameters that fuzzy comprehend evolution (FCE) would use to calculate the safety level of conflict points, which are divided into four levels, safety (level 1), relative safety (level II), less safety (level III), dangerous (level IV). After analyzing, the result of this study was indicated following the table 1 as below:

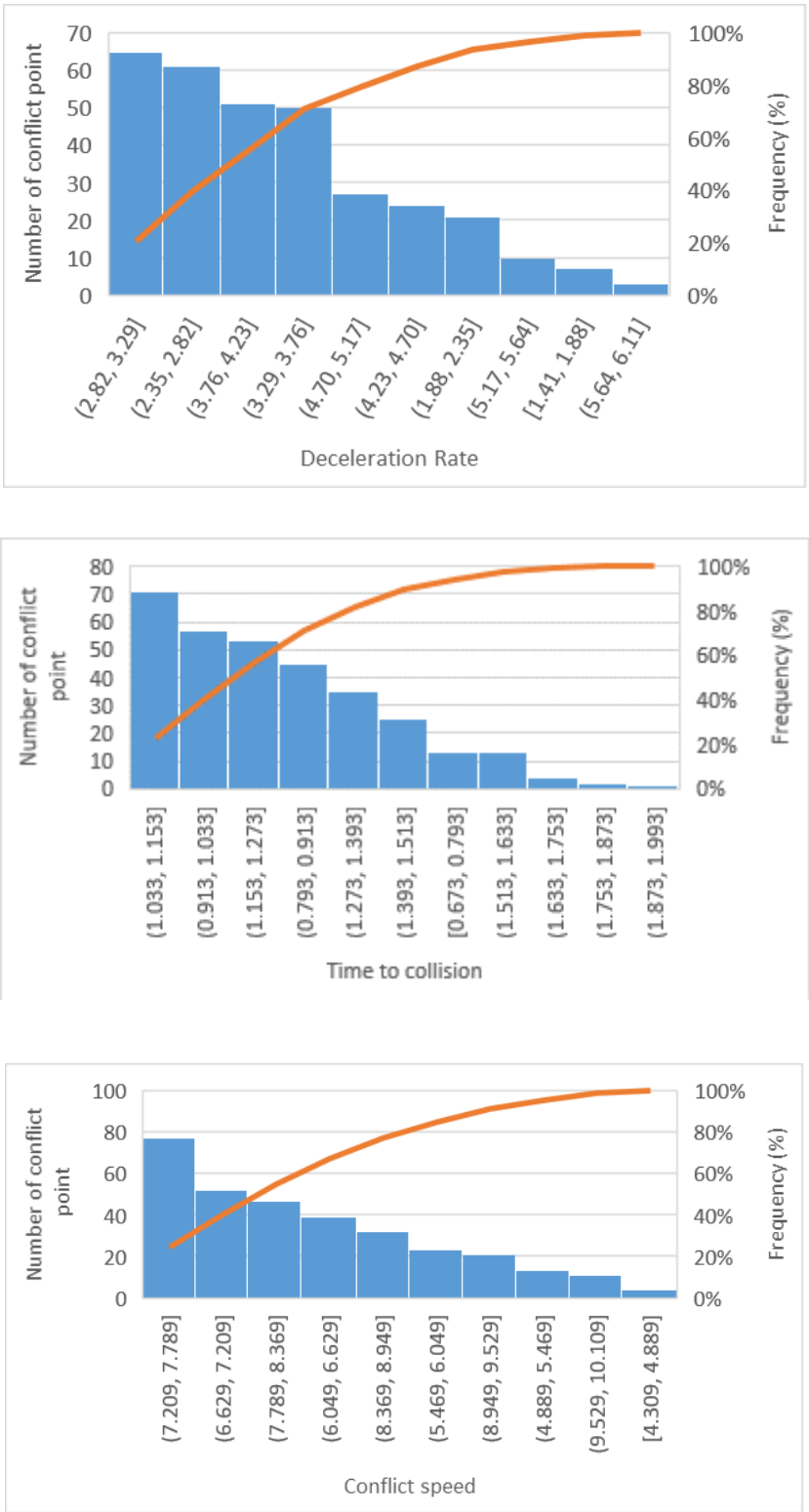


Figure 2. *Cumulative frequency curves of conflict parameters collected (DR,TTC,CS).*

Data was recorded by video processing are input parameters that fuzzy comprehend evolution (FCE) would use to calculate the safety level of conflict points, which are divided

into four levels, safety (level 1), relative safety (level II), less safety (level III), dangerous (level IV). After analyzing, the result of this study was indicated following the table 1 as below:

Table 1. *Data collected and analysed.*

STT	Conflict Speed (m/s)	Time to collision TTC (s)	deceleration rate DR (m/s^2)	Correlation coefficient of different safety levels				Level of safety
				I	II	III	IV	
1	6.68	1.41	2.36	0.84	0.16	0.00	0.00	I
2	8.35	1.26	3.29	0.10	0.42	0.25	0.22	II
3	7.53	1.52	2.48	0.67	0.08	0.25	0.00	I
4	6.31	0.84	3.74	0.30	0.03	0.40	0.27	III
5	8.91	1.05	4.23	0.00	0.00	0.36	0.64	IV
6	7.33	1.10	3.31	0.00	0.64	0.36	0.00	II
…	….	….	…..	….	….	….	….	….
319	6.485	0.976	3.32	0.2423	0.3738	0.2799	0.1040	II

Parameters of conflict points were showed in 3-dimensional space and classified in 4 regions correlation 4 levels of safety.

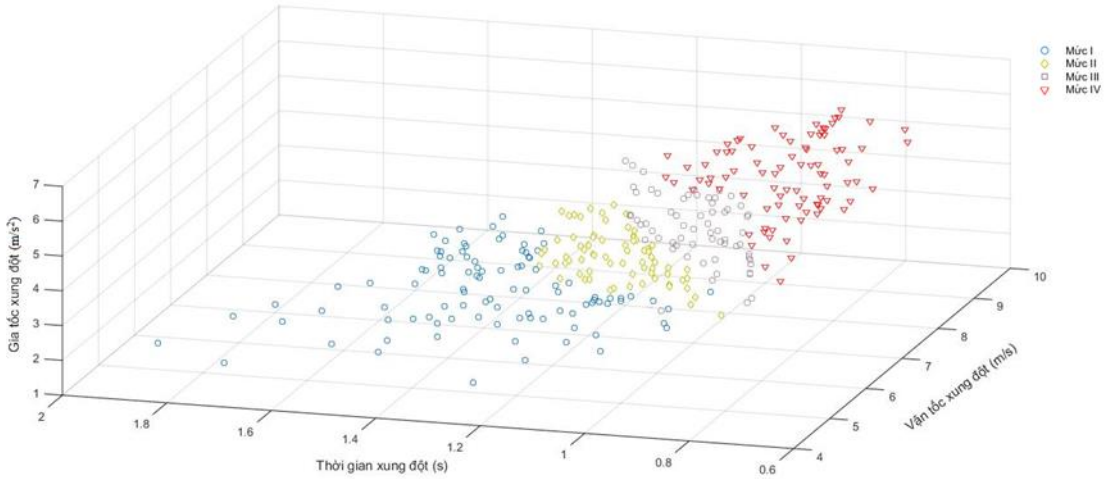


Figure 3. *Clustering follow by level of safety of conflict point.*

The heat map was, also, used to present the distribution and location of dangerous conflict points (level 4), different colors correspond to different numbers of conflict occur in the survey period. The red color expresses the areas having many dangerous conflicts, the yellow indicates less and there are some dangerous conflicts in green spaces.

Figure 4. *Heat maps demonstrated location and concentration of conflict point.*

At Nguyen Chanh - Mac Thai To intersection, the conflicts are evenly distributed in areas of the intersection, while Van Cao - Lieu Giai - Doi Can intersection, hazardous conflicts improve traffic light signal control. To solve this issue, a new control scheme was considered to replace the current plan, in which 3 phases were installed to prevent conflicts that occur between vehicles go straight an turn left. To assess the effectiveness of this option, a traffic scenario was created by PVT – Vissim, which is a traffic simulation software and SSAM3, which is a software used to perform safety analysis of traffic facilities.

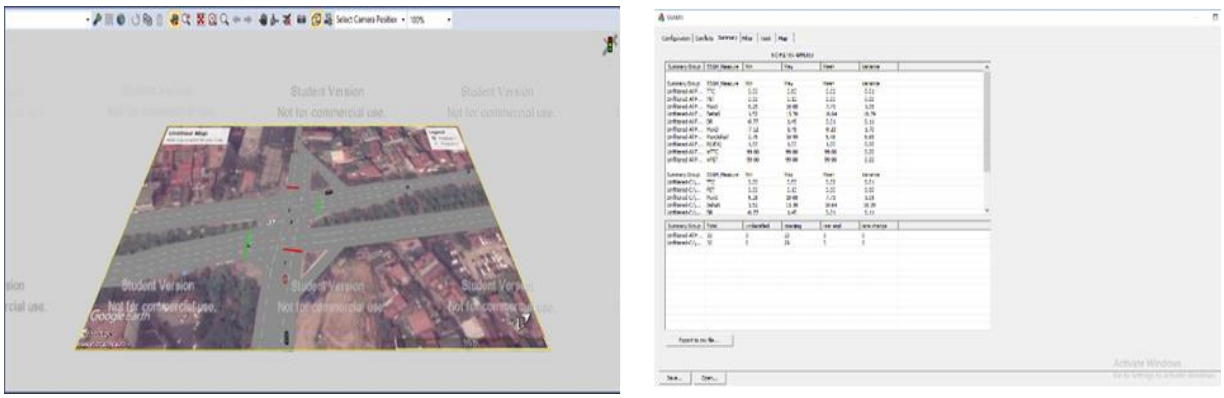


Figure 5. *Traffic scenario simulation by using PVT- Vissim software and window in SSAM3.*

The results indicated that the total of danger conflict point reduces when a new plan was implemented. In particular, after running simulation, there are 72 conflict cases when applying the old scheme and 38 situations that were collected in the proposed option, about 47.2% off. Many dangerous conflict points also decrease around 71.42%.

4. CONCLUSION

This paper presented the method of conflict analysis at an intersection using image processing technique and fuzzy comprehensive evaluation. The results of the research describe point conflicts and it's the location that provides the background of traffic at the intersection based on traffic conflict technique and thus will have an effective solution to solve some remaining problems. However, the study has not established a relative between conflict data and traffic accident, so it has not fully shown the role of conflict data set, it will

be done in the next study in the future.

REFERENCES

[1] Jutaek Oh, Eungcheol Kim, Myungseob Kim, Sangho Choo: Development of conflict techniques for left-turn and cross-traffic at protected left-turn signalized intersections. Safety Science 48 460–468, 2010.

[2] Perkins, R.S, and Harris: Traffic conflict characteristics-Accident potential at intersections. Highway Research Record No. 224 35-43, J.I. 1968.

[3] Hayward: Near misses as a measure of safety at urban intersections. Dissertation, The Pennsylvania State University, J.C. 1971.

[4] Allen, B.L., Shin, B.T., and Cooper. Analysis of traffic conflicts and collisions. Report No. TRR 667, Transportation Research Board, Washington, D.C, P.J. 1978.

[5] L. A. Zadeh, "Fuzzy sets," Information control, vol. 8, no. 3, pp. 338-353, 1965.

[6] Lai Zheng, Karim Ismail, and Xianghai Meng: Traffic conflict techniques for road safety analysis: open questions and some insights, NRC Research Press, 2014.

AUTOMATIC CRACK DETECTION FOR TUNNEL LINING CONCRETE USING IMAGE PROCESSING AND GENETIC ALGORITHM

Nguyen Kim Cuong[1,2*], Kei Kawamura[2], Hideaki Nakamura[2]

[1]*University of Mien trung civil engineering, Ha Huy Tap, Phu Yen, Vietnam.*

[2]*Yamaguchi University, Ube, Yamaguchi, Japan.*

[*]*Corresponding author: nguyenkimcuong@muce.edu.vn*

Abstract: Automatic crack detection is a main task in a crack map generation of the existing tunnel lining concrete inspection. This paper showed an automatic crack detection based on genetic algorithm (GA) to optimize the parameters of image processing techniques (IPTs). The experimental results of the crack detection for tunnel lining concrete surface images under various complex conditions validate the reasonable accuracy in practical application. Moreover, the proposed method is compared with our previous methods.

Keywords: Crack detection, Tunnel Lining, Image processing, Genetic algorithm

1. INTRODUCTION

Many concrete components of existing infrastructure systems such as bridges, and tunnels have suffered from various geologic, loading and environmental conditions cause to cracks which make influent to quality of operations. Therefore, the condition assessment of the existing tunnel lining is an important task not only for warming against deterioration but also for guaranteeing soon maintenance. Concrete cracks are important indicators reflecting the safety of tunnel lining. The automatic crack detection based on image data has been considered significantly due to accuracy, objectivity and timing inspection. This technique can be implemented using some of different image data captured from ultrasonic device, infrared and thermal device, laser scanning, and commonly digital cameras.

Image processing techniques consist of three approaches: edge detection, threshold technique there are main three approaches based on threshold technique (Ito et al. 2002 [1]; Miyamoto et al. [2]; Fujita el al [3]) and mathematical morphology (Sinha et al 2016 [4]). Machine learning algorithms (MLAs) are commonly used to decide the parameter value of IPTs. Therefore, they are applied to detect and classify concrete infrastructure surface cracks. In recently years, many automatic crack detection and classification methods based on a combination of IPTs and MLAs are implemented as decision tree (DT) (Kei et al.2013 [5]), support vector machine (SVM), k-clustering nearest neighbour (K-NN), and artificial neural network (ANN) (Zhang et al.2014 [6]).

Moreover, the parameter-optimization algorithms of IPTs such as genetic algorithm (GA), particle swarm optimization (PSO), artificial bee colony (ABC), and differential evolution (DE) are typically utilized. In recently years, Kawamura et al. 2003 [7] proposed interactive genetic algorithm is applied to adjust the image processing parameters. It is

developed to extract crack pattern effectively. Nishinkawa et al. 2012 [8] designed System for Automatic Construction of Image Filter (SACIF) based on a genetic program to detect cracks on concrete structures.

The main aim of this paper is to combine genetic algorithms (GA) adjusted optimized parameters and IPTs which are capable of detecting the cracks of the various complex surface images of the concrete infrastructures.

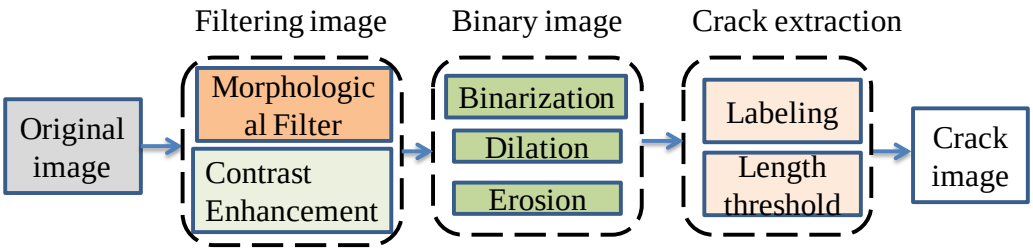


Figure 1. *Image processing techniques.*

2. PROPOSED METHOD

2.1. Image processing techniques

Figure 1 shows the image processing techniques that compose of three main parts. Namely, there are filtering image part, binary image part, and feature extraction part. Therein, filter image part comprises of morphological transform and contrast enhancement of the gray-scale image. The binary image part consists of binarization, and dilation-erosion transform of the binary image. To the end, geometric transform composes of labelling and linear degree.

2.1.1 Morphological filter

The morphology filter consists of opening transform of closing transform on gray-scale image with a predefined structuring element as the following equation:

$$M = \max([(G \oplus S)\Theta S]\Theta S] \oplus S], G) - G \quad (1)$$

Where *M* means the smooth image after morphological image processing, S is a structuring element, G is a gray-scale image converting from the corresponded original image. Additionally opening transform includes dilation of erosion. Closing is an inverse operator of the opening as the following Eqs. (2) and (3):

$$(G \oplus S)_{(x,y)} = \max_{u,v}(G_{(x-u,y-v)} + S_{(u,v)}) \quad (2)$$

$$(G\Theta S)_{(x,y)} = \min_{u,v}(G_{(x-u,y-v)} - S_{(u,v)}) \quad (3)$$

The purpose of this part is firstly to smooth image as well as eliminate noises and shading and then make the sharpness of crack edge pixels.

2.1.2. Structuring element design.

Because the shape of crack is irregular elongation, the shape of the structuring element of line-type is adopted. To retrieve fully crack information in the various directions,

combination line types of four directions into the structuring element such as {0°, 45°, 90°, 135°} or six directions {0°, 30°, 60°, 90°, 120°, 150°}. In this study, the former type is used to reduce computational volume.

2.2. Contrast enhancement

Surface images of civil infrastructure are the typical dark environmental condition, narrow contrast. The intensity of crack pixels and background pixels is not much difference. Therefore, a histogram of the raw image is equalized before binarization to sharpen crack pixel as the following equation:

$$q_k = \sum_{i=0}^{k} \frac{n_i}{n} \quad k = 0,1,2,..L-1$$
$$I_k(\text{out}) = (I_{max} - I_{min}) \times q_k(\text{in}) + I_{min} \tag{4}$$

Where, n_i: is number of pixels that have i^{th} gray-scale value. n: the number of pixels; k: the input gray-scale level. L: the maximum gray-scale level (255). q_k: acquired normalized histogram. I_{max} and I_{min} are respective to the maximum and minimum intensities of the original image. I_k(out) is the new gray-scale intensity of the input pixel having q_k(in).

2.2.1. Binarization

As shown in Eq. (5), the purpose of binarization is to segment gray-scale image into binary image. The binary image only has two values 255 (white pixel is background) and 0 (black pixel is crack or noise).

$$P(i) = \begin{cases} 255 & if I(x_i) > T \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

Where *T* is a threshold value of the binarization (binary), it is necessary to find out an optimal value. *P(i)* is the i^{th} pixel value after binarization step.

2.2.2. Dilation

The aim of a dilation operation is to connect crack fragments meanwhile noises are separated from the cracks. Therefore it results in the reduction of loss pixel. However, widths of crack pixels increase along with the crack shape. The size of structuring element of the dilation operation is considered as an adjusted parameter to optimized value by GA. In this paper, the shape of structuring element is predefined in "square" type. The rule of the dilation operation is if any pixel in the input neighborhood is "1", the output pixel is "1". Otherwise, the output pixel is "0".

2.2.3. Thinning

Thinning to a binary image is a morphology operator used to remove noise surrounding crack pixels. The purpose of this step is to prune branches from the crack shape as well as reduce noises after dilation operator is performed.

2.2.4. Labeling and length threshold

After thinning, the pixel values of the image are presented by the 0s or 255s. The automatic labeling step is to combine the 8^{th}-connection components as a single object with the same number.

Hence, a single object can be a crack object or a noise object. Therefore, the length threshold decides as if the object is removed following equation:

$$P'^{(i)} = \begin{cases} 255 \ if \ \dfrac{Max\left(R_{xi},\ R_{yi}\right)^2}{S_i} < \mathrm{T}' \\ 0 \qquad otherwise \end{cases} \tag{6}$$

Where T' is a threshold value of the i^{th} label length to distinguish between a crack object and a noise object. S_i is the total number of pixels of the i^{th} label. R_{xi}, and R_{yi} are the number of pixels of the i^{th} label in the horizontal and vertical directions, respectively.

The purpose of the first part is to make blurred images as well as eliminate noises and shading. The main aim of the binarization is to segment grayscale image into binary image depending on a threshold value. The purpose of dilation/erosion that is processed in the binary image part is to connect fragment images of crack meanwhile noises are separated from the cracks. In the end, the threshold defined in linear degree is to decide whether the single objects in the binary image are removed.

The productivity of the filtered image depends on the parameter value of structuring element size in dilation and erosion transform. In binarization, if the threshold value is too high, many crack pixels are lost. If threshold value is too low, more noise will occur. Such it is necessary to find out an optimum threshold value. Similarly, the parameter values of dilation/erosion and linear degree in the binary image part also affect to the quality of output image. Therefore, these parameters are adjusted to the optimized values based on GA.

2.3. Application of GA to the image processing parameters optimization

2.3.1 Represented chromosome design for solution candidates

The parameters of IPTs are combined together for creating an individual in a population. Next each individual is represented by a chromosome encoded to a binary string, as shown in figure 2. Namely, the size of structuring element (f) is assigned by 6 bits, the threshold value of binarization (b) is expressed by 8 bits, dilation transform parameter (d) is expressed by 4 bits, and the linear degree is expressed by 6 bits (l). Table 1 shows the parameter value range which design based on the preliminary experiments.

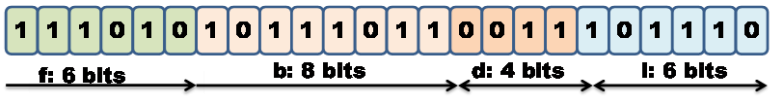


Figure 2. *A represented chromosome for solution candidates.*

Table 1*. Properties of parameter.*

Variable	Range	Steps	Bits
f	[1 127]	2	6
b	[0 255]	1	8
d	[1 31]	0.5	4
l	[0 32.5]	0.5	6

2.3.2. Genetic algorithm

Figure 3 indicates a sequence of GA including into the crucial three stages. Namely, they consist of the initial population generation, fitness evaluation of each individual in the current population, and evolution operation to create the next generation. Namely, the detailed steps are presented as the following three steps:

Step1. Generate initial population randomly

An initial population including 20 individuals was generated randomly with respect to 20 phenotypes to start fitness evaluation. To assess the fitness of the individual in the current population, an objective function to assess crack detection accuracy is defined as the Eq.(7). Loss and noise are computed based on comparison between the processed image and the target image shown in figure 5. As a result, the objective function (F) has to ensure the accuracy of extracted crack information with the minimum noises and losses as much as possible. The accuracy and the processing time can be considered as evaluation costs.

Step 2. Evolution operation

The evolution operation comprised of selection, crossover, and mutation is repeated until finding best solution. Each binary string encoded from the searching range of the parameter values has a corresponded fitness value. The probability of each string to be selected is proportional to its fitness value based on the Roultte wheel rotation randomly. The process is repeated for the second parent. Two elite members are kept forward to the next generation.

To improve quality of individual fitness, the crossover operation is used to create two new children from two selected parents with predefined probability. Crossover point is point laid on between 0 to the end of chromosome length. In this study, the single crossover point is selected. The part of the first parent chromosome that runs until the crossover point is spliced with the part of the second parent chromosome that includes, and runs after, the crossover point shown in figure 3. The whole new generation is selected in this manner. The mutation of bit strings ensue through bit flips at random positions. The purpose of the mutation operation is to create genotype diversification in the population in order to avoid local optimization leading to finding the best solution. Mutation point is chosen randomly. However, mutation rate is very small under 1% to avoid collapsing the genetic structure of the current population.

Step3. Stopping criterion

Evaluation of each individual meets the predefined maximum generation.

$$F=\frac{1}{\sqrt{w_1 f_1^{\,2}+w_2 f_2^{\,2}}},\ F\in[0\ 1]$$
$$f_1=\frac{\text{Loss}}{\text{Crack}};f_2=\frac{\text{Noise}}{\text{Back}} \quad (7)$$

Where Crack and Back are the number of black pixels and white pixels in the ground-truth image, respectively. In this paper, the weight parameters of the objective function $w_1=w_2=0.5$.

f_1, f_2 are loss rate and noise rate, respectively. F measures the accuracy of crack detection. F is larger value, the accuracy is higher.

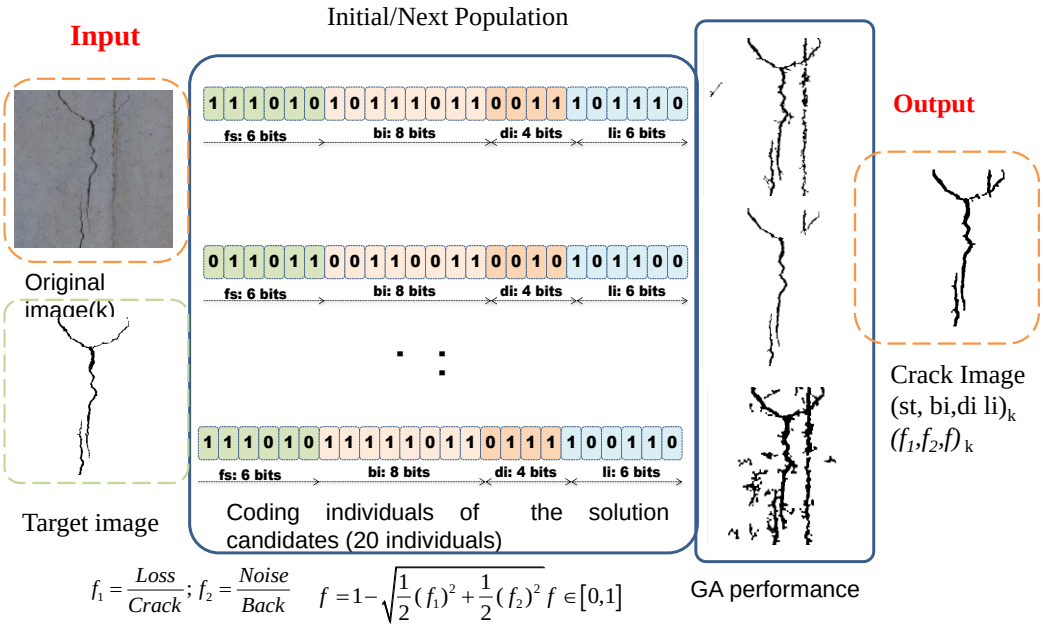
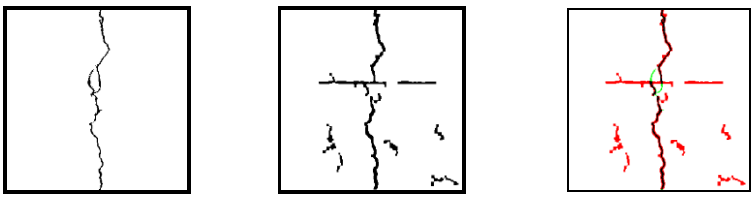


Figure 3. *Procedure of genetic algorithm.*



(a) (b) (c)

Figure 4. *Result of supervised learning algorithm with one-pixel accuracy. (a): Ground-truth image; (b): Processed image; (c): Compared image. A number of red pixels in the compared image are noise pixels annotated Noise. A number of green pixels in the compared image are loss pixels annotated Loss.*

3. EXPERIEMNT AND RESULTS

3.1 Sample test

The tunnel lining surface images were captured in various complex environmental conditions. Each inspection image of (3456x4608) pixels with 0.072 mm/pixel resolution was divided into 300x300 pixels to reduce computation time effectively. This paper used the following equation (8) to estimate the detectable minimum crack width.

$$w \geq 0.5 \times f \times r \tag{8}$$

Where *w* is the detectable crack width, *f* is the size of the structuring element, and *r* is the image resolution.

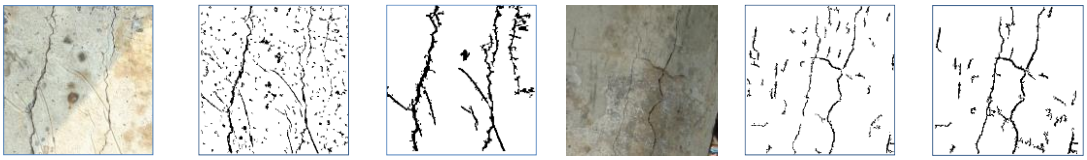
Therefore, the detectable minimum crack width of the proposed method would be 0.25 mm because the size of a structuring element is 7x7 pixels and the resolution of the original images in this paper is 0.072 mm/pixel. Figure 4 illustrates the crack detection result of a tested sample with one-pixel accuracy. Comparing to the ground-truth image, the result image still appears a number of loss pixels (green) and noise pixels (red).

3.2 The dilation and thinning transform experiment

Figure 5 shows the results of the experiment of the proposed method with and without dilation–thinning transform. As a result, without dilation-thinning transform, the accuracy of crack detection result decreases due to noises. Having dilation-thinning transform, the accuracy (*f)* of crack detection is improved significantly from 0.8575 to 0.9542 for Fig.5 (a), and from 0.8753 to 0.9595 for Fig.5 (b).

3.3 Performance evaluation

To validate the effectiveness of crack detection algorithms, the crack detection accuracy of the proposed method (denoted method 1) is compared with the one of our previous method called method 2 [9]. As can be seen in Fig.8, the accuracy value of crack detection of the method 1 is higher than the one of the method 2. However, the crack line of the method 1 is wider than the target image while the crack width of method 2 is similar to the one of the target image.



(a) (c) (e) (b) (d) (f)

Figure.5 *The effectiveness of the dilation and thinning transformation; (a)&(b): Original images; (c)&(d): Proposed method without dilation/thinning has the accuracy f=0.8575 & f=0.8753, respectively: (e) & (f): Proposed method with dilation/thinning has the accuracy f=0.9542 and f=0.9595.*

This reason is that the main aim of the method 1 is to remove the noises surrounding the crack line using the dilation-thinning transform. In contrast, the method 2 prefers to reduce the noise laying the crack line. These tendencies depend on image processing techniques in the methods. So far, the average accuracy of the crack detection using methods 1 and 2 are 0.9578 and 0.9202, respectively. It is concluded that application of dilation-thinning transform with the optimal size of the structuring element based on GA improves the crack

detection results extremely.

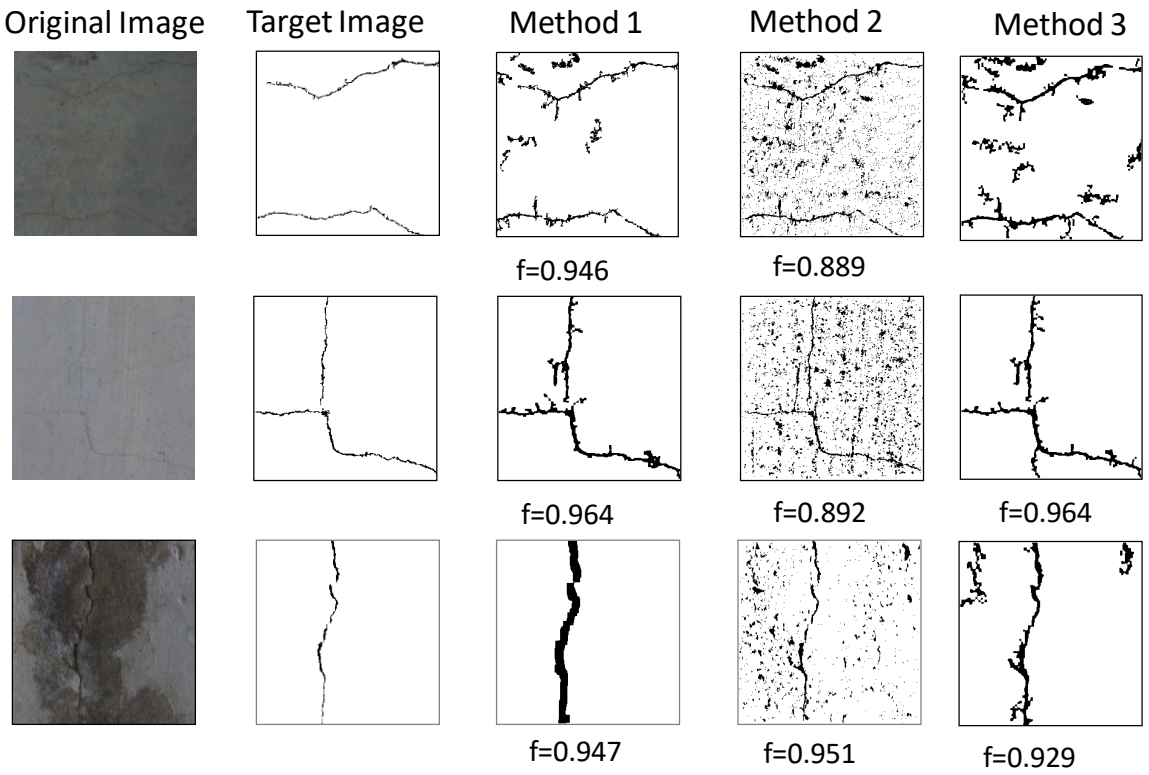


Figure.8 *The crack detection results of proposed method and our previous method. The first column is original images. The second column is the target images. The third column is the results of proposed method. The fourth column is the results of method 2. And, the fifth column is the results of method 3.*

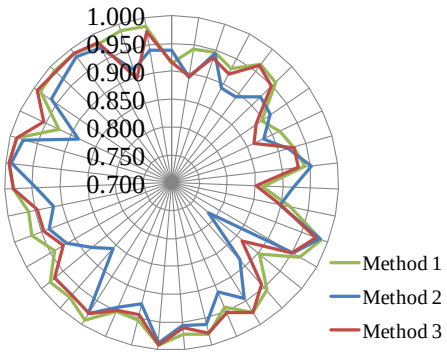


Figure 9. *Comparison of the accuracy of crack detection methods.*

Further, Figure.9 shows the results of the three methods using radar chart. Therein, method 3 is the proposed method using Otsu's threshold [10] instead of the global optimized threshold of the binarization by GA. As can be seen in Fig.9, the vertical axis is the accuracy of crack detection method, and the horizontal axis is number of sample tests included 40

images with various environmental conditions. As a result, the best crack detection accuracy belongs to method 1. The results of method 3 are better than method 2. There are not much difference between method 1 and method 3.

The sample tests in this experiment are not complete. However, they are effective enough to supply a sense of the ideas validated this work.

4. CONCLUSION

The experimental results of various low-contrast surface images of concrete tunnel lining under complex brightness conditions demonstrated that the effectiveness of the proposed method. However, the crack width results of the method 1 were larger than the ones of the target image. This reason was that the proposed method 1 prioritized to detect the thin cracks without losing the crack pixels. This method preferred to remove noises surrounding the crack line significantly. Moreover, the trade-off between the reduction rate of noise surrounding the crack line and the one laying the crack line depended on the type of the objective function. This problem would be solved in the future work.

ACKNOWLEDGEMENT

The research is supported by the financial support of JSPS KAKENHI Grant Number 15K0618, and scholarship of Ministry of Construction Vietnam.

REFERENCES

[1] Ito, A., Aoki, Y., and Hashimoto, S. (2002). “Accurate Extraction and Measurement of Fine Cracks fiom Concrete Block Surface Image” In IECON 02 Industrial Electronics Society, IEEE 2002 28th Annual Conference of the IEEE, 2202–2207.

[2] Miyamoto, A., Kono, M., and Bruhwiler, E., (2007). “Automatic crack recognition system for concrete structures using image processing approach”, Asian Journal of Information Technology, 6(5), 553–561.

[3] Fujita, Y., Mitani, Y., and Hamamoto,Y., (2006) “A method for crack detection on a concrete structure” ICPR2006: IEEE 18th Int.Conf, on Pattern Recognition, Vol.3, New York, 901-904.

[4] Sinha, K.S., Fieguth, P.W, “Automated detection of cracks in buried concrete pipe images”, Automation in Construction 15,2006, 58-72.

[5] Kawamura, K., Yoshino, K., Nakamura, H., Sato, Tarighat, A. (2013). A Valid Parameter Range Identification Method of a Digital Image Processing Algorithm for Concrete Surface Cracks Detection Using Genetic Algorithm and Decision Tree, Proc. Japan Soc. Civ. Eng. Vol.69.

[6] Zhang,W., Zhang, Z., (2014) “Automatic Crack Detection and Classification Method for Subway Tunnel Safety Monitoring”, Sensors, , 14, 19307-19328.

[7] Kawamura, K., Miyamoto, Nakamura, A. H. , Sato, R.: “Proposal of a crack pattern extraction method from digital images using an interactive genetic algorithm” Proc. Japan Soc. Civ. Eng.

Vol.742, 2003, pp. 115-131.

[8] Nishikawa, T., Yoshida, J., Sugiyama, T., Fujino, Y., "Concrete crack detection by multiple sequential image filtering", computer-Aided civil and infrastructure engineering 27, 2012, pp 29-47.

[9] Nguyen Kim, C., Kawamura,K., Tarighat, A., Matsumoto, J., and Shiozaki, M., (2016), "Development of semi-automatic crack detection software for concrete structures", 11th fib International PhD Symposium in Civil Engineering.

[10] Otsu, N., (1979) "A threshold selection method from gray-level histograms", IEEE Trans Syst Man Cybern 9:62-66.

ALASSO - IMPROVEMENT OF REAL ESTATE PRICING WITH AGGREGATION FEATURES TO LASSO REGRESSION

Pham Van Toan[1*], Nguyen Minh Tam[1], Ta Minh Thanh[2]

[1]*Sun Asterisk R&D Lab, 13F, Keangnam Hanoi Landmark Tower, Plot E6, Me Tri, Nam Tu Liem, Hanoi, Vietnam.*

[2]*Le Quy Don Technical University, 236 Hoang Quoc Viet, Co Nhue, Bac Tu Liem, Hanoi, Vietnam.*

[*]*Corresponding author: pham.van.toan@sun-asterisk.com*

Abstract: LASSO (stands for least absolute shrinkage and selection operator) is a regularization method that is applied to both linear and nonlinear regression to remedy over-fitting (high variance) in high-dimensional statistical models. In this article, we apply the LASSO method for variance reduction for the nonlinear regression model in the real estate evaluation problems. Even though only a handful number of properties are considered in this task – linear models normally lead to under-fitting (high bias) problems, making non-linear regression the model of choice. When we apply these types of models, the number of parameters is consequently large, results in a relatively higher model's capacity compared to linear ones. As the result, it is prone to over-fit the underlying distribution, leading to the need for regularization. In this task, the LASSO method becomes our choice for two main reasons: effective in both model's useful attributes (variable) selection and regularization. However, applying LASSO with non-linear regression comes with a demerit: sensitiveness with the choice of the penalization hyper-parameter. We proposed an ensemble method to overcome this shortcoming. This method has been evaluated in the real estate dataset collected in Montreal province, Canada, and Long Bien District in Ha Noi, yielding promising results.

Keywords: real estate pricing, nonlinear regression, linear regression, LASSO method, nonlinear regression shrinkage via LASSO.

1. INTRODUCTION

The real estate market is one of the most important markets in the modern economy. Usually, real estate is the largest asset a person ever owns. It is not only a place of residence but also able to serve as an investment. The more information about the real estate, the better decision making and risk management for both selling and buying parties. Although there are many other factors involved, the lack of information about the real estate can cause a significant loss for one party. Therefore, any source of information that allows a relatively more accurate real estate evaluation should be of great concern to all parties involved in the transaction (Mu et al., 2014).

Evaluating housing prices is a non-trivial task due to its nature of complexity and the

large finance transaction involved. Traditional real estate evaluation method includes: determining which factors of the properties should be considered in the pricing process; and expert knowledge in analyzing how each factor, alone or combined, influences the assessment. This expert knowledge, or pricing modeling methods, is commonly kept confidentially by private real estate companies. making the properties estimation less transparent and comprehensible to the buyer parties. This leads to our motivation of modeling housing prices by learning from data, to lessen the dependence of buyers on experts and strengthen the evaluation validity of sellers.

Determining the factors affecting property prices is impactful, especially on urban planning. Those factors can be utilized to identify and develop the types of infrastructure that meet the needs of buyers.

In addition to traditional methods such as those being used in Vietnam, the world has been studying and applying widely mathematical modeling for real estate evaluation. Since 1963, Bailey et al. Have proposed a method of determining the price of a property using linear regression (Bailey et al., 1963). By 1964, Alonso, a famous researcher in the real estate sector, claimed that the value of property heavily depends on its location. In 1966, in consumer theory, Lancaster introduced a model in which the value of an asset was dependent on several properties. By 1974, Rosen used the ideas of Bailey and Lancaster to come up with a mathematical model, the Hodenic model, evaluate properties' worth, and analyze their market equilibrium value. During the following times, many studies applying the Hodenic model to real estate price determination were carried out by researchers in different countries specifically the work on real estate market in Lasvegas, USA (Zhou et al. 2008); France (Christian et al., 2009), Poland (Król, 2015). In general, in the Hoderic model, the price function of a property depends on its features including location, nearby public utilities, and infrastructure facilities, house area, number of bedrooms, number of floors, v. v .. Models used to approximate the relationship between price and its independent variables can be as simple as the linear model or more complex as the exponential or logarithmic models.

There have been several pieces of research on building real estate pricing models using machine learning algorithms. One of the most interesting efforts is real estate valuation in Montreal (Noseworthy et al., 2014). The results of this paper are impressive and influence the selection of attributes in our data. The authors used linear regression combined with the LASSO method and K-nearest neighbor. These results guide us to investigate, survey, and build real estate pricing models with data collected from Long Bien district. We apply our proposed method on these 2 datasets to see its effectiveness in real estate pricing modeling.

Our main contributions are as follow:

• We propose the LASSO nonlinear regression for housing prices evaluation, results in competitive performance compared to the latest algorithms for real estate valuation in the province of Montreal (Noseworthy et al., 2014).

• We construct a new real estate dataset, collecting data from Long Bien district with a different set of features to the Montreal dataset. Using this dataset sheds light on cases on which our proposal might not work.

• The performance of our method is sensitive to the choice of regularization hyper-parameter. To boost up the overall generalization and weaken this sensitivity, we apply the assembling method, specifically the weighted ensemble.

2. RELATED WORKS

2.1. Linear regression

The linear model is commonly used in real estate price determination (Christian et al., 2009; Richard, 2009; and Bartik, 1987) due to its simplicity and explain ability. Linear regression defines a line or a plane across data points in features space. Assume that the price of the real estate is and the attributes affecting its price such as area, frontage width, the width of the access road, the legal status of the land, amenity of residential area (sanitary conditions, school conditions, health), distance toward the center, district, city are quantified and denoted by $x_1, x_2, ..., x_p$. We formalize the price function of real estate as a linear function according to the above variables, as follow:

$$y = f(x_1, x_2, \ldots, x_p) = w_0 + \sum_{k=1}^{p} w_k x_k$$

Through data surveys, we collected a dataset containing n data points {(yi, xi)}, I = 1, 2, … n. Our objective is to specify coefficients $w_i, i = 0,1,2,...,n$ such that the square of the error, which is determined bellow, is minimal:

$$\min\left\{\frac{1}{2n}\sum_{i=1}^{n}\left(w_0 + \sum_{k=1}^{p} w_k x_k^i - y^i\right)^2\right\}$$

This is a convex optimization problem that can be solved with a closed form solution or iterative method such as gradient descent. Consequently, based on the magnitude of the weights, we can determine which parameters as well as which factors heavily influence housing evaluation.

2.2. LASSO nonlinear model

Despite its simplicity, the linear model suffers from high bias (under-fitting complex conditional distribution $y|x$) (Brunauer, et al., 2013). In this section, we propose the square root approximation of the price function as a function of the square roots of the variables.

$$\sqrt{y} = w_0 + \sum_{k=1}^{p} w_k x_k + \sum_{k=2}^{p}\sum_{l=k}^{k-1} w_{kl}\sqrt{x_k}\sqrt{x_l}$$

The real estate price function is, therefore, determined by the regression function:

$$y = f(_1, x_2, \ldots, x_p) = \left(w_0 + \sum_{k=1}^{p} w_k x_k + \sum_{k=2}^{p}\sum_{l=k}^{k-1} w_{kl}\sqrt{x_k}\sqrt{x_l}\right)^2 \quad (1)$$

Our objective becomes:

$$\min\left\{\frac{1}{2n}\sum_{i=1}^{n}\left(w_0+\sum_{k=1}^{p}w_k x_k^i+\sum_{k=2}^{p}\sum_{l=1}^{k-1}w_{kl}\sqrt{x_k^i}\sqrt{x_l^i}-\sqrt{y^i}\right)^2\right\}$$

Higher capacity than its linear version is both merit and demerit of this model as it over-fits datasets without sufficiently large sample sizes (Hastie et al., 2009). To remedy the situation of low bias and high variance, we apply the regularization method. We opt for LASSO for two reasons: its efficiency in estimates' variance reduction and interpretability as a parameter subset selection method (Hastie et al., 2009)). The LASSO method finds the coefficients w_k, w_{kl} by solving the optimization problem:

The LASSO problem is often rewritten in a Lagrange form:

$$\min\left\{\frac{1}{2n}\sum_{i=1}^{n}\left(w_0+\sum_{k=1}^{p}w_k x_k^i+\sum_{k=2}^{p}\sum_{l=1}^{k-1}w_{kl}\sqrt{x_k^i}\sqrt{x_l^i}-\sqrt{y^i}\right)^2+\lambda\left(\sum_{k=1}^{p}|w_k|+\sum_{k=2}^{p}\sum_{l=1}^{k-1}|w_{kl}|\right)\right\} \quad (3)$$

with $\lambda \geq 0$. Due to the Lagrange duality, there is an equivalence between the constrained optimization problem (2) and the Lagrange form (3).

2.3. Ensemble methods

The ensemble method is widely used for generalization error reduction by combining several weak models to form a strong one (both high bias and low variance) (Breiman, 1994). This combination can be performed by averaging the predictions of weak models. Ensemble methods vary according to different ways to construct weak models. Those models can be dissimilar by architectures, training methods, training datasets, or only initialization and hyper-parameter settings (Bengio, 2015). Overall, "model averaging is an extremely powerful and reliable method for reducing generalization error" (Bengio, 2015). With that in mind, we introduce a variant of model averaging to reduce the variance of our algorithm, which is further explained in the Proposed Method section.

3. DATASETS

The objective of this paper is to build a model of real estate valuation in Long Bien district based on its set of properties. We want to explore how the price of a home is affected by infrastructure ò the district. This is a task that requires solving a regression problem to determine the dependence between an asset's price and its properties. So far in Vietnam, there is a lack of in-depth statistical study on this issue.

3.1. Long Bien District Real Estate Database

To examine the proposed algorithms and models, we use the data set collected from the Long Bien district on 487 real estates. Their attributes include transaction price, area of land, location of land divided according to regulations of the Ministry of Natural Resources and Environment, width access to the house, the convenience of the entrance to the house, the distance to the city center, the distance to the nearest school, the quality assessment of the school, the assessment of the quality of medical services, the assessment of the status of the red number, distance to the nearest market, distance to the district center, frontage width of parcel, total floor area of the house, house features, etc.

3.2. Montreal Province Real Estate Database

In this paper, we also evaluate the performance of algorithms on a real estate data set collected in Montreal province. This is an internationally published data set. This sample data set includes standard descriptions of each home as well as the number of facilities within a 3 km radius of that home. All property data is retrieved from online real estate agencies such as DuProprio (duproprio.com) and Royal Lepage (royallepage.ca). These sites provide all the necessary descriptions for the properties of the home.

Of the 9717 data samples collected, those with unusually high selling prices (greater than $ 75,000,000) are discarded as outliers that could cause a large margin of error while applying a regression algorithm. (Noseworthy et al., 2014). The main properties that affect the price of the house such as the type of apartment, the area of residence, the surrounding apartments, the number of rooms, the geographic location, as well as the number of surrounding infrastructure such as police station, fire station, memorial, shopping mall etc.

Missing data are common in this dataset. Missing properties affect the valuation of the property. Noseworthy et al., 2014 offer three approaches to solving the missing information problem: removing records with loss properties, predicting lost value with the maximization expectation. (Moon, 1996) and replace the lost value with the mean of the attributes. The authors have shown that this method of removing missing data is most effective in building a pricing model. At that time the data set was reduced to only 2289 records. In this paper, the reduced data set will be used to evaluate the performance of the learning algorithms.

4. PROPOSED METHOD

The LASSO method and linear regression were applied effectively to real estate data set in Montreal province. However, the hypothesis of real estate prices following the linear model is not satisfactory (Król, 2014). Besides, LASSO linear regression is provided as a solution for a high-dimensional sparse-data regression problem (the number of attributes is large compared to the number of records). However, with only a few dozen features, we are not necessarily facing this problem (Noseworthy, 2014). In this paper, we choose a nonlinear regression model for real estate valuation, combined with LASSO regularization for variance reduction. Besides, as our proposal's performance is sensitive to the choice of regularization hyper-parameter, we introduce an aggregation method based on the ensemble learning principle, to combine weak (inaccurate) LASSO regression functions into a strong (more accurate) regression function. This results in a better generalization of the algorithm. The regression model is trained on the training data set and finally evaluated on the test data set. 5-folds cross-validation is applied for hyper-parameter choices. Our algorithm is described in the below sub-sections.

4.1 Ensemble strategy with LASSO nonlinear regression aggregation

The LASSO nonlinear regression aggregation algorithm applies several nonlinear regression models combined with the LASSO method as described above. However, the model's error fluctuates according to the choice of λ parameter. Therefore, in this paper, we overcome that disadvantage by assembling (corresponding to the difference λ). The algorithm consists of the following steps:

• Step 1: Find the w_k^0, w_{kl}^0 parameters from equation (3) corresponds to the initialization value $\lambda_0 = 0$, estimate the mean absolute error e^0 of the training data

• Step 2: Calculate $\lambda_m = \lambda_0 + m\nabla\lambda$, and find w_k^m, w_{kl}^m parameters from equation (3) corresponds to value $\lambda = \lambda_m$, estimate the mean absolute error e^m of the training data ($\nabla\lambda = 0.005$)

• Repeat step 2 with *m = 1, 2,…* util $e^m \geq e_0 + e^{\nabla}$ ($e^{\nabla} = 5000$), then in the last step we get *m = M*. These LASSO nonlinear regression models are aggregated to form a combination of nonlinear regression:

$$w_k = \frac{1}{M+1}\sum_{m=0}^{M} w_k^m, w_{kl} = \frac{1}{M+1}\sum_{m=0}^{M} w_{kl}^m$$

These coefficients are used to construct the final regression model for real estate valuation, the property price function given by formula (1). This method not only automatically performs significant coefficients selection, but also reduces the variance, improving the generalization of the algorithm.

5. EXPERIMENTS RESULT

The performance of the linear regression models, nonlinear regression with and without the LASSO, and the LASSO nonlinear regression aggregation were compared on the real estate data sets of Montreal province. Noseworthy et al., 2014 pointed out that linear regression has equivalent performance as LASSO linear regression and the authors also point out that these are suitable property valuation methods, giving good results on the data set. Data collected in the province of Montreal. With the experimental results shown below, we can see that the nonlinear regression aggregate gives relative accuracy about 2% more accurate than the above methods (average prices of real estate in Montreal). collected is $ 312380).

5.1. Linear regression and LASSO linear regression

Table 1: *Results of average absolute error (based on test data, Montreal province) corresponding to the values of the corrected parameters λ of LASSO.*

LASSO linear regression	$\lambda = 0$	$\lambda = 1.0$	$\lambda = 5.0$	$\lambda = 10$	$\lambda = 100$	$\lambda = 1000$
Error (mAE)	*46677*	*46676*	*46668*	*46654*	*46557*	*47383*

Table 1 shows the detailed results of the mean absolute error of linear regression method LASSO over the values $\lambda = 0,1,5,10,100,1000$. With $\lambda = 0$, the LASSO linear regression becomes the linear regression. We can see that the mean of error is relatively stable unless λ receive a very large value. The best performance of LASSO linear regression on the Montreal province real estate

dataset is corresponding $\lambda = 100$, it yields an average absolute error of \$46557 .

Table 2 gives the detailed results of the mean error of the LASSO nonlinear regression that was specifically identified in Section 3 of the Montreal prefectural property dataset. With $\lambda = 0$, this model reduces to the nonlinear regression model described in detail in Section 3. We can see that the mean absolute error of the nonlinear model is quite small for the training data (\$31749) but quite large for test data (\$52828). But the absolute average error of the LASSO nonlinear regression fluctuates much when running through the values $\lambda = 1,5,10,100,1000$. This is due to a large number of parameters (over 780) compared to the linear model.

Table 2: *The results of the mean absolute error corresponding to the values of the normalized parameters of the LASSO nonlinear regression on the training and test data set, Montreal province.*

LASSO Nonlinear Regression	$\lambda = 0$	$\lambda = 1.0$	$\lambda = 5.0$	$\lambda = 10$	$\lambda = 100$	$\lambda = 1000$
Training error	*31749*	*40036*	*43652*	*47840*	*80028*	*80028*
Testing error	*52828*	*43164*	*46502*	*51686*	*86664*	*86664*

We applied the LASSO nonlinear regression combined method for the Montreal province real estate dataset. We initialize the normalized parameter value $\lambda_0 = 0$ and $\nabla\lambda = 0.005$, and the absolute error mean difference threshold is $e^{\nabla} = 5000$. The mean absolute error of the LASSO nonlinear regression method is \$40250, that is 12.88% lower than the result reported by Noseworthy et al., 2014.

We also evaluate the performance of the LASSO nonlinear regression method with land price data collected by us in the Long Bien district. Primary raw data includes 50 properties and the real estate prices. However, this data contains many missing values. We exclude the uninformed 23 attributes, corresponding to it has 352 records containing full information of 23 selected attributes. The results are somewhat disappointing when the relative error of the LASSO nonlinear regression method is only achieved on 18.96% test data.

6. CONCLUSION

The results of real estate pricing in the Long Bien district are not as expected. The work (Noseworthy et al., 2014) initially gives us hope that we can achieve promising results when applying our method on Long Bien dataset. The use of different sets of attributes may be the reason why the mean absolute error obtained in our experiment cannot be compared with the results obtained on the real estate data for the province of Montreal. However, the results cannot be compared directly because the attributes in the data set for Montreal province and Long Bien district are inherently different. Furthermore, the Long Bien District property data set, after removing many attributes, may not contain all the information needed for real estate valuation. Furthermore, the majority of data is collected from real estate owners. Intuitively, if the houses are close to each other, if they have similar properties, their cost should be similar, but in this data set sometimes it is not. The fact is that homeowners all tend to judge their homes very subjectively.

However, there are several encouraging results from our experiments. Most importantly, the successful construction of a real estate pricing algorithm on the Montreal province is an

internationally published real estate data set. Our experiment has shown that the LASSO nonlinear regression aggregation algorithm is better than the latest algorithms for the Montreal province real estate data set (Noseworthy et al., 2014) with a lower error of 12.88%. For linear models, their performance degradation can be explained by the non-linearity of the property price function. This is a reasonable hypothesis due to the complexity of the real estate market.

ACKNOWLEDGMENT

This work is partially supported bySun-Asterisk Inc. We would like to thank our colleagues at Sun-Asterisk Inc for their advice and expertise. Without their support, this experiment would not have been accomplished.

REFERENCES

[1] Alonso W. (1964). Location and Land Use. Harvard University Press, Cambridge, Massachusets.

[2] Bailey M. J., Muth R. F., Nourse H. O. (1963). A regression method for real estate price index construction, Journal of the American Statistical Association, section 58, page 933 - 942.

[3] Bartik T. J. (1987). The estimation of demand parameters in hedonic price models, Journal of Political Economy, section 95, no 11, page 81 - 88.

[4] Brunauer W., Lang S., Umlauf N. (2013). Modeling House Prices using Multilevel Structured Additive Regression, Statistical Modelling, section 13, no 2, page 95 - 123.

[5] Christian G., Laferrère A. (2009). Managing hedonic housing price indexes: The French experience, Journal of Housing Economics, section 18, page 206 - 213.

[6] Hastie T., Tibshirani R., Friedman J. (2009). The Elements of Statistical Learning Data Mining, Inference, and Prediction, Springer.

[7] Hastie T., Tibshirani R., Wainwright M. (2015). Statistical Learning with Sparsity, The Lasso and Generalizations, CRC Press.

[8] Król A. (2015). Application of Hedonic Methods in Modelling Real Estate Prices in Poland, Data Science, Learning by Latent Structures, and Knowledge Discovery, page 501 -511.

[9] Lancaster K. J. (1966). A new approach to consumer theory, Journal of political economy, section 74, page 132 - 157.

[10] Moon T. K. (1996). The expectation-maximization algorithm, Signal processing magazine, IEEE, section 13, no 6, page 47 - 60.

[11] Mu J., Wu F., and Zhang A. (2014). Housing Value Forecasting Based on Machine Learning Methods, Abstract and Applied Analysis, section 2014, 7 page. doi:10.1155/2014/648047

[12] Noseworthy M., Schiazza B. L. (2014). Montreal Real Estate Pricing, Technical Report, McGill University, Website: http://rl.cs.mcgill.ca/comp598/fall2014/comp598_submission 89.pdf.

[13] Peter E. (1998). Improved Price Indexes for Real Estate: Measuring the Course of Swedish Housing Prices, Journal of urban economics, section 44, page 71 - 96.

[14] Richard J. C. (2009). The Hedonic Pricing Model Applied to the Housing Market of the City of Savannah and Its Savannah Historic Landmark District, The Review of Regional Studies,

section 39, no 1, page 9 - 22.

[15] Rosen S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition, Journal of Political Economy, section 82, no 1, page 35 - 55.

[16] Tibshirani, R. (1996). Regression Shrinkage and Selection via the LASSO, Journal of the Royal Statistical Society, section 58, no 1, page 267 - 288.

[17] Zhou W. X., Sornette D. (2008). Analysis of the real estate market in Las Vegas: Bubble, seasonal patterns, and prediction of the CSW indices, Physica A, section 387, page 243 - 260.

INTENT DOMAIN IDENTIFICATION USING DEEP LEARNING MODEL

Luong Thai Le [1*]

[1]*University of Transport and Communications, No 3 Cau Giay Street, Hanoi, Vietnam.*

[*]*Corresponding author: luongthaile80@utc.edu.vn*

Abstract: Nowadays, people are willing to share their feeling, activities and even their plan about what they intend to do on online social media (e.g., buy an apartment, borrow amount of money, travel to somewhere, etc.). Understanding those users intentions is a crucial need in a lot of different business areas, such as real estate, tourism, e-commerce and online marketing. However fully understanding user intents is a complicated and challenging process which includes three major stages: user intent filtering, intent domain identification, and intent parsing and extraction. In this paper, we propose the use of convolution neural network (CNN) model to classify intent-holding posts into one of several categories/domains. The proposed method has been evaluated on a medium-sized collections of intention posts in Vietnamese, and the empirical evaluation has shown promising results with an average accuracy of 77%.

Keywords: intention mining, user intent identification, domain classification, social media text understanding, text classification.

1. INTRODUCTION

Today, many Internet users commonly share their feeling, daily activities, and even their intention on online social media channels like Facebook and Twitter. For example, one user may post *"I intend to book a summer holiday trip at Ha Long bay for my 4 people family"* or *"We are looking for an apartment near the downtown center"* on a discussion forum or on his/her own Facebook wall. Those posts are called *"intention posts"* because they carry user intents to do something in the near future. If enterprises know and understand exactly what online users are planning to do from these kind of posts, they can easily locate a large number of potential customers relevant to their business domain.

Being aware of this important trend, many previous researches focused on the understanding of user intents behind their online activities like web search [1, 8, 9, 12, 14] or computer/mobile interactions [15]. Most of these studies attempted to guess or determine the user *implicit* intents behind their search queries and browsing behaviors. Understanding search intent helps improving the quality of web search significantly. *Explicit* intent, on the other hand, is a directly or explicitly written statement by a user about what he or she plans to do. According to Bratman (1987), intent or intention is a mental state that represents a commitment to carrying out an action or actions in the future [2]. As more and more users are willing to share their intents explicitly on the web, we have an opportunity to access to an

invaluable source of knowledge about a huge number of online users or probably potential customers. However, there have been few previous studies really focusing on analyzing and identifying user *explicit* inptents from their posts or comments on forums or social networks. This is explainable. In spite of its huge potential for application, the identification of user explicit intents is actually a natural language understanding problem which is inherently a hard research direction in natural language processing.

However, it can be solved gradually by narrowing its scope and domain and dividing into multiple steps. In our previous work [11], we have defined a three–stage process for user intent identification that includes three main tasks: (1) user intent filtering, (2) intent domain identification, and (3) intent parsing and extraction. The first task has been addressed in our previous work as normal classification problems [11]. In the scope of this paper, we focus on solving the second stage *(intent domain identification)* that helps to determine what an intention post is about. We consider this problem as a classification task, that is, each intention post is classified into a most suitable domain/category. This classification task is actually a text categorization problem where the input texts are short and quite ambiguous. There are several challenges in this task. First, an intention post commonly contains several sentences and it is sometimes very hard to determine the real domain of a post. For example, a post like *"I am going to buy a seven {seater car for traveling at weekend".* This intention is about *"buying a car",* however it can also be classified into tourism because it contains the word *"traveling".* The second challenge is that intention posts on online social media are very diverse. The number of specific domains is usually very large as users can share their intention about anything. It is very hard to perform a classification task with large number of classes. Therefore, we only classify intention posts into one of 12 major domains like *electronic device, fashion & accessory, finance, food service, furnishing & grocery, travel & hotel, real estate, job & education, transportation, health & beauty, sport & entertainment,* and *pet & tree.*

In this paper, we propose to use CNN model [16] to deal with this problem. We have conducted experiments with real data crawled automatically from four well-known discussion forums and social networks. We built a medium-sized labeled data set of text posts in Vietnamese for evaluation. And we achieved promising results.

2. PREVIOUS WORK

User intent understanding can be defined in different ways for different application domains. In this section, we will review several studies on understanding user goals or intents that are more or less related to our work. To the best of our knowledge, there is no one studying intention mining for text documents until 2013. Most of them are for web search where they focuseased on intent identification for search queries. Rose and Levinson (2004), Jansen et al. (2004), Kathuria et al. (2010), they all tried to understand the user intent from web search queries by classifying the queries into three major categories: *informational, navigational,* or *transactional* [8,10]. Baeza-Yates et al. (2006) presented a framework for the identification of the user's interests based on the analysis of query logs from web search engines. They first attempted to find the user goals and then mapping those queries into the categories: *informational, not-informational,* and *ambiguous,* and eighteen categories of topic

to classify the queries into. Almost all categories are based on The Open Directory Project1 [17]. Azin Ashkan et al. (2009) used the features of query based, content of search result pages and ad clickthrough to classify queries into two dimensions: *commercial, non-commercial* and *navigational, informational* [1].

The following studies are most relevant to our work. Chen et al. (2013) claimed that their solution is the first one that try to identify user intents in discussion forum posts. They proposed a new transfer learning method to classify the posts into two classes: *intent posts* (positive class) and *non-intent posts* (negative class) [4]. This work is most similar to our previous work that solves the first stage (user intent filtering) [11] in the user intent understanding process. But there is still a little difference between their work and ours: while they only consider purchase intents in four domains *cellphone*, *electronic*, *camera*, *tivi*, our work handles a lot of intent types, such as purchase, sell, hire, rent, borrow etc. and in a wide range of domains. Similarly, Gupta et al. (2014) attempted to identify only purchase intent from social post by categorizing the posts into two classes namely *PI* and *non-PI*. This has been done by extracting features at two different levels of text granularity, that are word and phrase based features and grammatical dependency based features [5]. More relevant to our work, Wang et al. (2015) attempted to mine user intents in Twitter by classifying tweets into six categories *food and drink, travel, career and education, goods and services, event and activities,* and *trifle* [18].

3. INTENTION DOMAIN IDENTIFICATION

3.1. Domain identification as a stage of intention understanding process

As we proposed in our previous paper [11], the process of analyzing and understanding user intents includes three major stages that be described below.

- *Stage 1 - User intent filtering*: this phase helps to filter text posts on online social media channels (blogs, forums, online social networks) to determine which posts
  contain user intents and which do not. Posts carrying user intents will be forwarded to the next stage bellow. This is actually a binary classification problem and has been solved in our previous work [11].
- *Stage 2 - Intent domain identification*: given a text post containing a user intent, this phase will analyze and identify the domain of the intent. This is the main problem we are aiming at to solve in this paper. In our work, an intent can be classified into one of the following categories: *electronic device, fashion & accessory, finance, food service, furnishing & grocery, travel & hotel, real estate, job & education, transportation, health & beauty, sport & entertainment, pet & tree*. This is actually a multi-class classification for short and ambiguous texts.
- *Stage 3 - Intent parsing and extraction*: given a text post containing an intent and its domain category, this phase will parse, analyze, and extract all concrete information (i.e., properties) of the intent. For example, if an intent is about tourism, its properties may be *destination, transportation, time-period, number of people*, etc.

In our previous work, we aimed to solve the *user intent filtering* phase by proposing a binary classification model to filter the intent posts from online Vietnamese social media texts. In this paper, we focus to solve the second phase - *intentdomain identification*, that determine the most suitable domain for each intent. We propose the set of twelve intent domains and CNN model for our experiments.

3.2. The set of intention domain

Building the set of intent domains turns out to be a difficult task. We had to discuss several times among data annotators to agree on a most suitable partitioning for intent posts. Each partition is considered as an intent domain. It means we want to make sure that if an intent post belongs to one domain, it cannot be assigned to any other domains.

After carefully analyzing the set of data and referring to several reference web sites in Vietnam (*https://www.consumerbarometer.com/about, https://www.chotot.com*), we decided to divide the intent posts into thirteen domains as shown in Table 1. Since an intent post maybe appear in the middle of a long conversation that the clear intention was mentioned at the beginning, it is difficult to identify its domain if only based on the post. For example, a user may write "*I'm going to buy the same one too*" or "*ship 1 kg for me at this weekend*". It is so difficult to understand the exact intent domain for these posts although we know that the posts carry purchase intents. Moreover, there are some posts simultaneously express more than one intent. For example, a post like "*I want to buy a second|hand eating chair for my baby. By the way, I'm looking for an extra job to have more income*" may be categorized in two different domains (*furnishing & grocery* and *job & education*). It will make the work more complicated. In the scope of this paper, we do not consider these sorts of posts. It means we only consider classifying posts that contain intention belong to only one clear domain.

The chart in Fig. 1 shows the percentage of each intent domain. The data were crawled from several famous discussion forums in Vietnam and from Facebook, such as *http://www.webtretho.com/forum, https://www.lamchame.com/forum, http://sotaychame.com /diendan.html, https://www.chotot.com,...* This can be considered the distribution of intention domains for Vietnamese users intention through online social media. As we can see, the domain *job & education* has the highest frequency, less frequent domains are *real estate, furnishing & grocery, transportation* and *fashion & accessory*.

3.3. Intent domain classification model

In this paper we proposed to use the CNN model to solve the task of intent domain identification. This model is inspired by the CNN model that Y.Kim proposed in his research for text classification. In this section, we will describe our proposed model including 4 layer as bellow.

- **Input word embedding layer**: Each word in the training data is encoded by a 300-element vector and generated in a random normal distribution. After that, an intention post of n words will be encoded into n interconnected vectors. This vector will be adjusted during training the model.
- **Convolution layer:** The model uses 100 filter windows with 3 sizes of 3, 4, and 5 words respectively. The activation function used to train the model is the ReLU. So, each filter window will create a feature vector with corresponding size.

Tabel 1. *Intent domain descriptions and examples.*

Intent Domain	**Descriptons / Example**	**#**
Electronic Device	I want to liquidate the old refrigerator… I have an old breast pumps want to sell…	546
Fashion & Accessory	I was presented a pair of leather shoes, but they do not fit me, so I want to sell them ... Is there any mum here know a nice fashion clothes store, please show me, I need to buy a new dress…	586
Finance	I urgently need to borow a huge amount of money ... I'm looking for someone who can make capital contribution ...	314
Food Service	This weekend, I have some nice bacon, who want to buy, please order with me... I'm looking for a restaurant to celebrate my son's birthday...	424
Furnishing & Grocery	Is there any mom here want to liquidate a dinning chair for kid, I need one... I'm finding a brand new wardrop...	699
Health & Beauty	I'm going to buy a pressure cuff for my mother ... I really want to have a nose-lift performed ...	322
Job & Education	I have a pressing need of finding a domestic helper ... I'm looking for an English class of communication for my 12-year-old child ...	1296
Other	I need a smart accounting software ... I'm looking for a souvenir for my girlfriend ...	228
Pet & Tree	I need to sell my dog because I have no time to take care for him ...	385
Real estate	I'm going to buy an apartment ...the price is about 1.5 million (Vietnam dong) ... For hire, shop premises with frontages on two streets ...	750
Sport & Entertainment	I want to find a swimming class for my son ... I have a pair of tickets for Le Quyen liveshow this Saturday, want to resell ...	456
Transportation	I'm looking for a new 7-seater car to replace my old one ... I have a redundant air ticket to Sai Gon, need to resell ...	649
Travel & Hotel	I want to book a travel tour for 3 people to Nha Trang ...	354

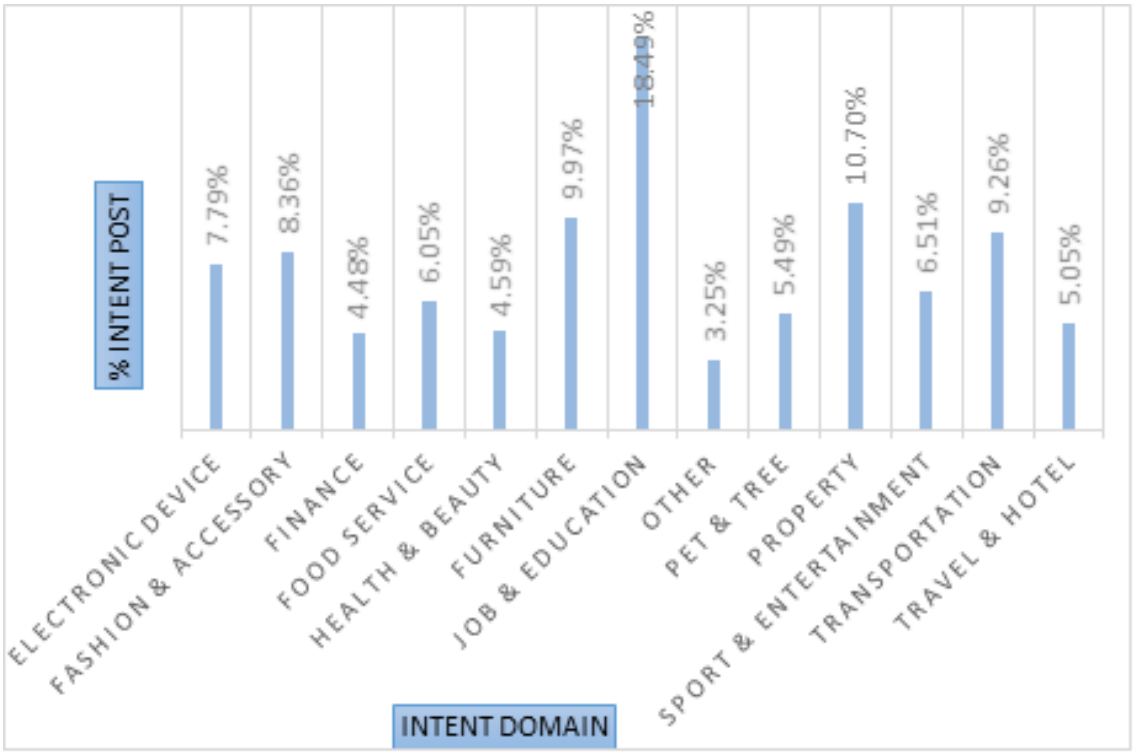


Figure 1. *The statistic of intent posts from our real life data.*

- **Maxpooling layer:** From each feature vector created from the convolution layer above, this layer extracts a feature component with the maximum value. These selected feature elements will be connected together to create the final unique feature vector. This vector will be moved to the next layer of normalization.
- **Softmax layer:** The final normalization layer ensures that the output value is a vector with each probability component belonging to one of 13 out-put classes. The class achieves the highest probability value will be chose as the out-put intent domain.

In the CNN model, we used the Adam optimization method with the standard $L_2 = 1$. For all experiments, we use *precision*, *recall* and *F*1-*score* as the evaluation measures.

4. EXPERIMENTS AND RESULTS

4.1. Experimental data and design

We collected a medium-sized collection of intent posts from some famous discussion forums in Viet Nam, such as Webtretho, Lamchame, Chotot, Sotaychame and facebook. After removing all unexplicit instances that we mentioned above in the subsection 4.1, our data collection consists 7009 intent posts. A group of students were asked to label the data into thirteen domains based on a given guideline and the agreement among them. Some examples for each intent domain can be seen in the Table 1. And the Fig. 1 also gives the statistic of the intent domains with the highest and lowest intent post numbers respectively belong to the domain *Job & Education* and *Other*.

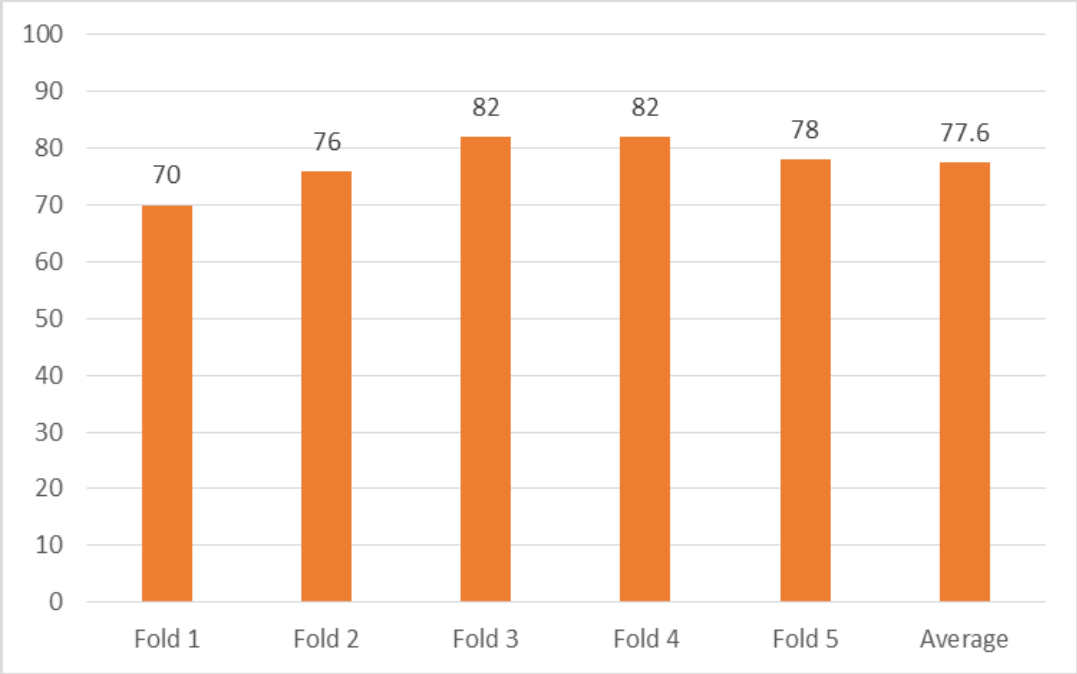


Figure 2*. F1-score for 5-fold cross validation accuracy.*

The labeled data collection were then divided randomly into five parts, in which four parts are used for training and the remainder is used for testing. The experiments were then done using 5-fold cross validation and the results will be reported in the next subsection.

4.2. Experimental results and discussions

As conducting 5-fold cross validation experiment, we shows the accuracy (i.e., micro-averaged *F*1 - score) of each fold and the average over the five folds in the Fig. 2. We can see that the results are quite stable over the four folds, and the average of F1-score is about 77.6%. This is a quite promissing high result because the size of the data set is not big enough for a deep learning model. This shows that the proposed classification model can work well on this data set. Moreover, with this deep learning model, we do not need to spend time and effort to build the set of technical features as doing with the traditional machine learning models.

We then show the accuracy for each class – intent domain in the result of fold 3 which achieved the highest F1-score in the Fig. 3. We can easily see that the domain **other** always has the lowest accuracy. This is understandable because of two reasons: (1) the number of intent posts belonging to the *other* class is smallest (accounts for only 3.25% of our total labeled data); (2) the *other* class contains miscellaneous intentions (as been mentioned in Table 1) that we cannot place them in any of the twelve intent domains. Thus it makes the CNN model difficult to learn the feature itself for this class. However, except the *other* class, we can find that the results are quite stable over the remaining twelve domains even though the number of intent posts for these domains are unequal. For example, *job & education* class has the number of intent posts be about three times as many as that of *travel & hotel* class, but as we can see in Table 2 that the *F1* measure of these two class are almost the same. This

shows that the classification models can work well on this data set.

Table 2. *Accuracy for each intent domain in fold 3.*

Intent Domain	**Precision**	**Recall**	**F1**
Electronic Device	86%	58%	70%
Fashion & Accessory	76%	81%	78%
Finance	96%	87%	92%
Food Service	85%	88%	87%
Furnishing & Grocery	68%	91%	78%
Health & Beauty	82%	61%	70%
Job & Education	89%	97%	93%
Other	80%	9%	16%
Pet & Tree	88%	91%	89%
Real estate	86%	93%	90%
Sport & Entertainment	69%	59%	64%
Transportation	78%	92%	85%
Travel & Hotel	89%	80%	84%

5. CONCLUSION

In this paper, we have presented the problem of domain identification for intention posts and proposed our solution to this problem. We considered this problem as a multi-class classification task. To evaluate, we crawled real posts from online social media, filtering posts containing user intents and performing domain annotation. By this way, we have built a medium-sized labeled dataset for conducting the experiments and proposed a set of twelve intent domains for classification. Then we have built our classification models with CNN method. Although the size of experimental data is not big enough for a deep learning model, all experiments achieved significantly high results (about 77.6% of accuracy on average). In the future work, we will crawl more data to improve the result.

REFERENCES

[1] A. Ashkan, C. L. Clarke, E. Agichtein, and Q. Guo. Classifying and characterizing query intent. In Proceedings of The 31th European Conference on Information Retrieval (ECIR), pp.578{586, 2009.

[2] Bratman, M.: Intention, plans, and practical reason. Harvard University Press, 1987.

[3] L. Chen. Understanding and exploiting user intent in community question answering. PhD Dissertation, Birkbeck University of London, 2014.

[4] Z. Chen, B. Liu, M. Hsu, M. Castellanos, and R. Ghosh. Identifying intention posts in

discussion forums. In Proceedings of The The North American Chapter of the Association for Computational Linguistics (NAACL), pp.1041-1050, Atlanta, Georgia, June 2013.

[5] V. Gupta, D. Kedia, D. Varshney, H. Jhamtani, and S. Karwa. Identifying purchase intent from

[6] social posts. In Proceedings of The AAAI Conference on Artificial Intelligence, 2014.

[7] B. Hollerit, M. Kroll, and M. Strohmaier. Towards linking buyers and sellers: detecting commercial intent on Twitter. In Proceedings of The World Wide Web Conference (WWW), pp.629-632, 2013.

[8] J. Hu, G. Wang, F. Lochovsky, J. T. Sun, and Z.Chen. Understanding user's query intent with Wikipedia. In Proceeings of the World Wide Web Conference (WWW), pp.471 - 480, 2009.

[9] B. J. Jansen, D. L. Booth, and A. Spink. Determining the user intent of web search engine queries. In Proceedings of The World Wide Web Conference (WWW), pp.1149 - 1150, 2007.

[10] Jethava, V., Liliana, C.B., Ricardo, B.Y.: Scalable multi–dimensional user intent identification using tree structured distributions. In: The ACM SIGIR, 2011.

[11] A. Kathuria, B. J. Jansen, C. Hafernik, and A. Spink. Classifying the user intent of web queries using k-means clustering. Internet Research, Vol.20, No.5, pp.563 - 581, 2010.

[12] T.L. Luong, T.H. Tran, Q.T. Truong, T. M.N. Truong, T.T. Phi, and X.H. Phan, Learning to filter user explicit intents in online Vietnamese social media texts. In Proceedings of The Asian Conference on Intelligent Information and Database Systems. (ACIIDS), pp.13 - 24, 2016.

[13] X. Li: Understanding the semantic structure of noun phrase queries. In: ACL 2010.

[14] C. Gonzalez-Caro. Supervised identification of the user intent of web search queries. The PhD Dissertation, Pompeu Farbra University, Spain, 2011.

[15] Y. Shen, J.Yan, S.Yan, L. Ji, N. Liu, Z.Chen: Sparse hidden–dynamic conditional random fields for user intent understanding. In: The WWW, 2011.

[16] K. Church, B. Smyth: Understanding the intent behind mobile information needs. In: The 14th IUI, 2009.

[17] Y. Kim, Convolutional neural networks for sentence classification." arXiv preprint arXiv: 1408.5882, 2014.

[18] R. Baeza-Yates, L. Calderon-Benavides, C. Gonzalez-Caro. The intention behind web queries. String Processing and Information Retrieval, pp.98 - 109, 2006.

[19] J. Wang, G. Cong, W. X. Zhao, X. Li. Mining user intents in Twitter: a semi{supervised approach to inferring intent categories for tweets. In Proceedings of The AAAI Conference on Artificial Intelligence, pp.339 - 345, 2015.

NHÀ XUẤT BẢN GIAO THÔNG VẬN TẢI

80B Trần Hưng Đạo - Hoàn Kiếm - Hà Nội
Điện thoại: 024. 39423345 * Fax: 024. 38224784
Website: www.nxbgtvt.vn * Email: nxbgtvt@fpt.vn

Chịu trách nhiệm xuất bản: Nguyễn Minh Nhật
Biên tập: Nguyễn Ngọc Sâm
Thiết kế bìa: Phân hiệu Trường Đại học GTVT tại TP. HCM
Trình bày: Phân hiệu Trường Đại học GTVT tại TP. HCM

Đối tác liên kết xuất bản: Trường Đại học Giao thông vận tải
Địa chỉ: Số 3 phố Cầu Giấy, Láng Thượng, Đống Đa, Hà Nội

SÁCH KHÔNG BÁN

In 300 cuốn, khổ 20,5 x 29,5 cm tại Công ty TNHH MTV DV TM Ngọc Linh.
Địa chỉ: C1/1 Lê Văn Việt, KP2, P. Tăng Nhơn Phú A, Quận 9, TP. Hồ Chí Minh
Số xác nhận đăng ký xuất bản: 903-2020/CXBIPH/1-27/GTVT.
Mã số sách tiêu chuẩn quốc tế - ISBN: 978-604-76-2102-6.
Quyết định xuất bản số: 159 LK/QĐ-XBGT ngày 16/11/2020.
In xong và nộp lưu chiểu năm 2020.