Journal of Knowledge Information Technology and Systems

ISSN 1975-7700

http://www.kkits.or.kr

Peak Hour Identification for Traffic Congestion Based on IoT Environments

Vasanth Ragu, Saraswathi Sivamani, Myeongbae Lee, Hyunwook Cho, Yongyun Cho, Jangwoo Park, Changsun Shin*

Department of Information & Communication Engineering, Sunchon National University, Suncheon-si, Republic of Korea.

ABSTRACT

This study deals with the analysis of traffic congestion and the peak hour identification by using Kalman Filter and Ensemble Model. There are different types of traffic congestion, Roadway Traffic congestion, Airways Traffic congestion, Network Traffic congestion, and so on. This study focuses on Roadway Traffic congestion. The peak hour identification is essential to prevent roadway traffic congestion. In roadway traffic congestion, there are two categories in traffic data, namely Roadside Equipment (RSE) data and Video Detection System (VDS) data. Both data were collected from RSE devices and VDS devices, which are located in roadways signals, toll plaza, private sectors, and etc. In traffic data, it may contain error values. So, this paper applies the Kalman Filter for the purpose of removing the error values or inaccurate values and providing the cleaned Traffic data. The suggested study also uses the Ensemble Model to average the traffic data at corresponding hours easily to analyse the traffic data. To identify peak hour, it defines four different models by considering numbers, average times, and average speeds of vehicles. With the suggested method, the perfect peak hour in the traffic data can be easily and exactly obtained. In the tests and results, this paper showed the detailed process of peak hour identification in traffic congestion.

KEYWORDS: Ensemble model, Kalman filter, RSE, VDS, Traffic data, and IoT.

ARTICLE INFO: Received 10 April 2018, Revised 4 June 2018, Accepted 8 June 2018.

*Corresponding author is with the Department of Information & Communication Engineering, Sunchon National University, Suncheon-si, Republic of Korea, 57922. E-mail address: csshin@sunchon.ac.kr

1. Introduction

At present, traffic congestion occurs everywhere in the world and it affects every citizen of the country. There are different types of traffic congestion, those are airways traffic congestion, roadways traffic congestion, network traffic congestion, and etc. This study deals with roadways traffic congestion. Without control the traffic congestion then it affects our national economic and development of the country. So, traffic congestion is an issue of high significance in the world. Already, people follow the solutions such as traffic signals, traffic police, route diversion, and etc. But, we need to find out the peak hour and busy hours of traffic congestion. Therefore, this study focuses on peak hour identification of traffic congestion.

In this study, we analyze the traffic congestion and identify the peak hour and busy hours in roadways traffic data. In traffic data, there are two categories, namely Roadside Equipment (RSE) data and Video Detection System (VDS) data. These data are collected from roadside equipment devices and video detection systems devices which are located in highways, toll plaza, high traffic areas, private sectors, and etc. It is mainly fixed in national highways and toll plaza. Traffic data (RSE and VDS data) consist of the number of vehicles (counts), an average speed of the vehicle (km/hr), and an average travel time of the vehicle (seconds), but VDS didn't have the average travel time of the vehicle. The Kalman filter is used to remove the error values or inaccuracy values with the help of a dynamic linear model. Next, we implement ensemble model which is used for purpose of averaging the data. In traffic data, it is important to analysis every hour of traffic congestion. So, we must average every hour of traffic data because of the need to identify busy hours and peak hour in traffic congestion. Therefore, this is the reason to adapt Ensemble model to analysis traffic congestion.

Afterward, we create different methods of peak hour identification in this study. These methods based on some conditions. If the conditions are satisfied, then it is simple to identify peak hour in traffic congestion. In result and discussion, the results of RSE and VDS plots explain the clear view of peak hour identification in traffic congestion. Section 2 describes related works, Kalman filter and Ensemble Model. Section 3 describes materials and methods of traffic congestion with specifically RSE & VDS and also describes methods of peak hour identification.

2. Related Works

Richard Arnott and Kenneth Small, examined the detailed view of the economics of traffic congestion[1]. Richard Arnott investigates the information to drivers reduce traffic congestion[2]. Yuki Sugiyama, Minoru Fukui, and et al, were analyzed traffic jams without bottlenecks and their experimental evidence for the physical mechanism of the formation of a jam[3]. Sandra Rosen bloom had been published peak period traffic congestion: a state of the art analysis and evaluation of effective solutions[4]. Christoph

Sommer and et al, examined the road traffic simulation for improved inter-vehicle communication and bi-directionally coupled network[5]. Venkataraman Shankar and et al, investigates the effect of roadway geometrics and explores the frequency of occurrence of highway accidents[6]. Yibing Wang and Markos Papageorgiou had been published the real-time freeway traffic state estimation based on extended Kalman filter a general approach[7]. L. Fei, and et al, examined the analysis of traffic congestion induced by the work zone[8]. Changzhi Bian and et al had been published the evaluation, classification, and influential factors analysis of traffic congestion in chinese cities using the online map data[9]. Eleonora D'Andrea and Francesco Marcelloni investigate the real-time detection of road traffic congestion and incidents from GPS trace analysis[10]. Kai Shi and et al, examined the detrended cross-correlation analysis of urban traffic congestion and NO2 concentrations in chengdu[11]. Li Wei and Dai Hong-ying have published the real-time road congestion detection based on image texture analysis[12]. Quddus and et al, were analysed the road traffic congestion and crash severity: econometric analysis using ordered response models[13]. Chao Wang and et al, examined the impact of traffic congestion on road accidents: A spatial analysis of the M25 motorway in England [14]. Zuduo Zheng and et al, had been published the impact of traffic oscillations on freeway crash occurrences[15]. David Mfinanga and Emmanuel Fungo, examined the impact of incidents on traffic congestion in Dar es Salaam City[16]. Saman Roshandel and et al, published the impact of real-time traffic characteristics on freeway crash occurrence: Systematic review and meta-analysis[17]. Based on these references, this study deals analysis of traffic congestion and identification of peak hour with the help of Kalman Filter and Ensemble Model.

2.1 Kalman Filter

Kalman filter is an algorithm that uses a series of measurements observed over time, containing statistical noise and other inaccuracies, and produces estimates of unknown variables that tend to be more accurate than those based on a single measurement alone, by using Bayesian inference and estimating a joint probability distribution over the variables for each time frame[18]. Kalman filtering is also called as linear quadratic estimation (LQE). In this study, Kalman filter is used for purpose of removing unwanted data which means removing inaccuracies or error values of traffic data and produces an output of more accuracies of traffic data.

2.2 Ensemble Model

Ensemble model is defined as simultaneously running more than two processes related but different analytical models and then synthesizing only one result into a single score[19]. In other words, an average is taken across the entire data of all possible time series that is called ensemble model in time series[20].

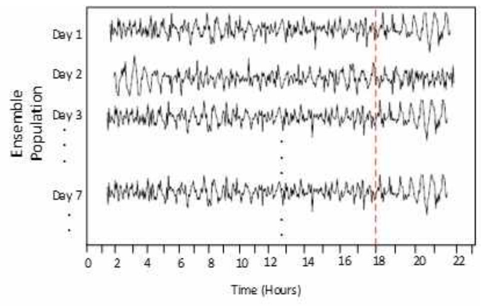

Figure 1. Diagram for Ensemble of Time Series.

In figure 1, the x-axis is time (hours from 0 to 23), the y-axis is the value of everyday ensemble population growth and the vertical red line shows for an average taken across the entire population at the particular time 18th hour (also possible to apply series of time) which is called ensemble population. Similarly, we can easily apply in traffic data and we will get the output of series of hours.

3. Materials and Methods

Commonly, RSE and VDS devices placed in many places, example: highway traffic, toll plaza, private sectors, and etc. These devices are used for counting number of vehicles, capturing the speed of vehicles, capturing travel time of vehicles at particular route. When the vehicle crosses the RSE device, it must have a hi-pass. Generally, hi-pass is used to pay highway tolls neither stop the car nor hand over the cash. Hi-pass needs onboard unit (OBU) and hi-pass card. If the vehicle has hi-pass when crossing the RSE device, then the device automatically connected to the vehicle and also easily fetching the information of a vehicle that can be stored in the information storage database. If the vehicle doesn't have a hi-pass, then the RSE device doesn't work on highways.

The VDS stands for video detection system, which is used for capturing each and every vehicle when crossing the VDS place and store the information in the storage database. In VDS device, there is no need to use OBU and hi-pass card, because it's only capturing the vehicles. It takes more storage capacity when compared with RSE device. Sometimes, both RSE and VDS devices are placed at the same point on highways. Figure 2 diagram shows for RSE and VDS placed at the same point in highway traffic.

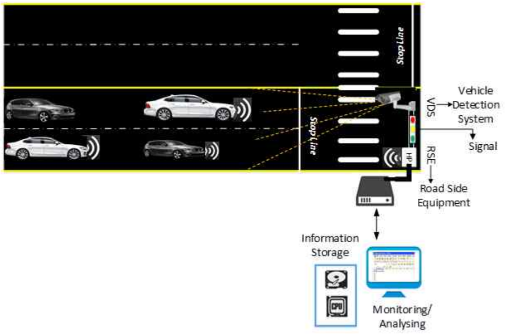

Figure 2. Diagram shows for RSE and VDS placed in Highway Traffic.

In traffic data, both RSE and VDS data consist of Section ID, Start Section Name, End Section Name, Date, Hours, Number of Vehicle (counts), Average Timing (travel time of vehicle), Average Speed (speed of the vehicle), and etc. But, the average time is not available in VDS data. We need only a few categories of traffic data, they are Date,

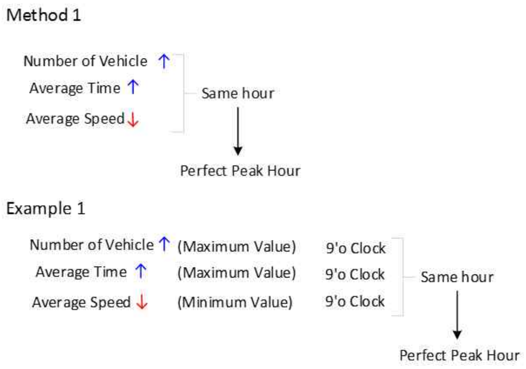

Figure 3. Diagram shows for Method 1 of Perfect Peak Hour Identification.

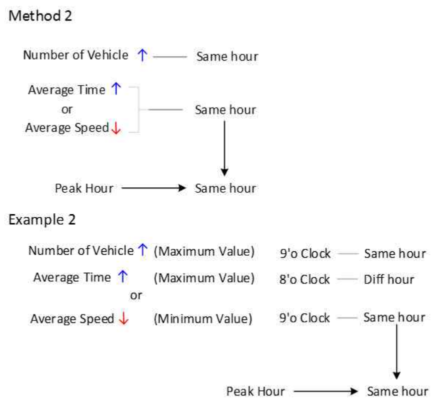

Figure 4. Diagram shows for Method 2 of Peak Hour Identification.

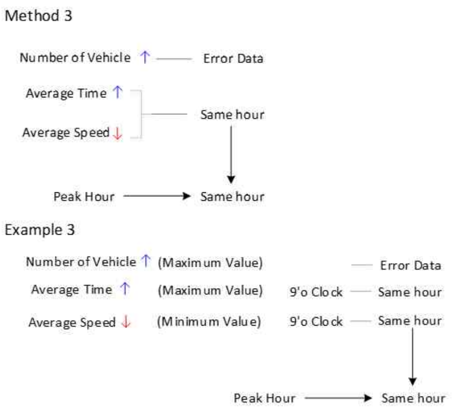

Figure 5. Diagram shows for Method 3 of Peak Hour Identification.

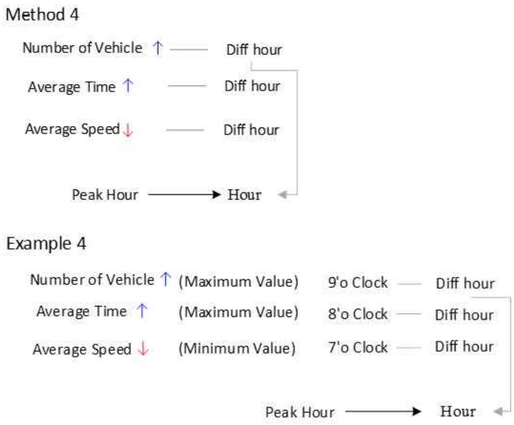

Figure 6. Diagram shows for Method 4 of Peak Hour Identification.

Hours, Number of Vehicle, Average Time, and Average Speed. So, these categories are used to identify the roadway traffic congestion. In these categories (number of vehicles, average time, and average speed) of traffic data, sometimes it has error values or inaccuracy data.

Using Kalman filter, to filter the error (inaccuracy) values and produce an output of accuracy values by R language. Implement ensemble model in traffic data, that is, mean value of traffic data (no. of the vehicle, average time, and average speed) with corresponding hours (from 0 to 23 hours) and then identify the busy hours and peak hour in traffic congestion.

We have four methods identify peak hour in traffic congestion. They are Method 1 Method 2, Method 3, and Method 4. These methods depend on certain conditions, and if the condition is satisfied then it is perfect peak hour. Otherwise, unable to identify the peak hour in traffic congestions.

There are three reasons for traffic congestion, they are a number of vehicle increases, average time (travel time) increases and average speed (speed of the vehicle) decreases. Those are all performed at the same time, then it is called perfect peak hour in traffic congestion. Figure 3 shows for Method 1 of perfect peak hour identification in traffic data. Example 1 shows for a number of vehicle increase at 9 o'clock, average time increases at 9 o'clock, and average speed decreases at 9 o'clock. So, at morning 9 o'clock is a perfect peak hour in example 1 of Method 1. In these methods, a number of vehicles are the first priority and others are next priority. In Method 2 of example 2, the number of the vehicle increases at 9 o'clock, average time increases at 8 o'clock, and average speed decreases at 9 o'clock. So, this type of situation we choose a first priority (number of vehicles) with a corresponding hour matched with other priority (either average time or average speed) then 9 o'clock is peak hour in example 2 of Method 2. Figure 4 diagram shows for Method2 of peak hour identification.

In Method 3, the first priority of the number of vehicles sometimes we get an error data or missing values, that time unable to get proper results. So, we just omit that first priority of category (number of vehicles) and taking the remaining categories (average time and average speed). In example 3, the number of vehicles has error data, average time increases at 9 o'clock and average speed decreases at 9 o'clock. So, in this method 3 for example 9 o'clock is the perfect peak hour in traffic congestion. Figure 5 diagram shows for Method 3 of peak hour identification.

Finally the Method 4 as completely different from others because of example 4, the number of vehicle the increases at 9 o'clock, the average time increases at 8 o'clock, and average speed decreases at 7 o'clock. Each category has different hours, so this type of situation we can choose the first priority of values, i.e. 9 o'clock is peak hour in traffic congestion. Figure 6 diagram shows for Method 4 of peak hour identification.

4. Result and Discussion

In this study, used only 1-year similar traffic data, because it's difficult to get real data. In general, x-axis has time (hours) from 0 to 23 and y-axis has series of the number which is used for a number of the vehicle, average time and average speed. Figure 7 shows up the result of Route 1 of RSE 1-year (Nov 2015 – Oct 2016) traffic data. Route 1 is the name of a particular route in traffic data. The green dotted-dashed line refers to average travel time (average time) of the vehicle and the range of travel seconds increased from morning 6 o'clock to evening 6 o'clock. The blue dotted-dashed line refers to the average of the count (number of the vehicle) and it is increased from morning 8 o'clock to evening 6 o'clock. The pink dotted-dashed line refers to the average of average speed of the vehicle(km/hr) and its decreased from morning 8 o'clock to evening 6 o'clock. In RSE traffic data, at morning 8 o'clock to evening 6 o'clock are increased numbers of the vehicle, average time, and decreased average speed of the vehicle. So, these hours are the busy hour on Route 1 of traffic data. The vertical red line refers to the maximum value of a number of the vehicle, the maximum value of average time and a minimum value of the average speed of the vehicle. So, it is satisfied Method 1 of peak hour identification. Therefore, 18th hour (evening 6 o'clock) is a perfect peak hour on Route 1 of RSE traffic data.

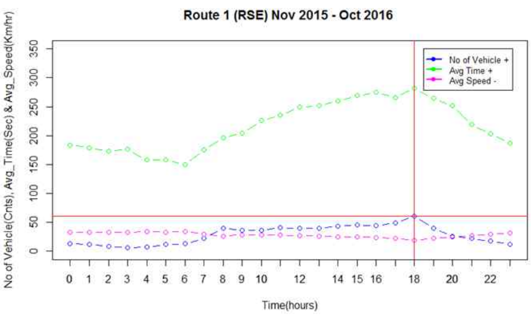

Figure 7. The result of RSE (Route 1) 1-Year Traffic Data.

Figure 8 shows the result of VDS (Route 1) 1-year (Nov 2015 – Oct 2016) traffic data. In VDS traffic data, two important categories, they are the number of vehicles and average speed of the vehicle.

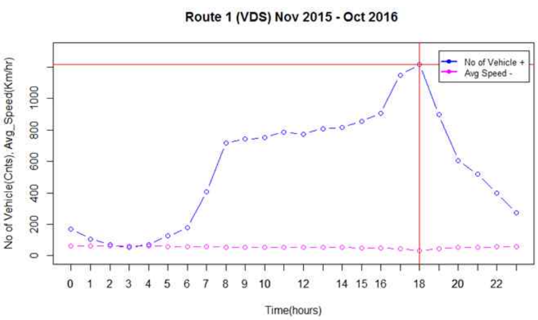

Figure 8. The result of VDS (Route 1) 1-Year Traffic Data.

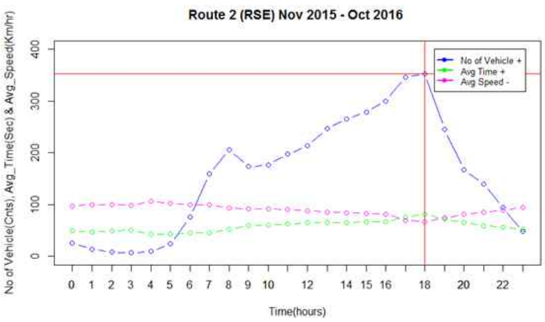

Figure 9. The result of RSE (Route 2) 1-Year Traffic Data.

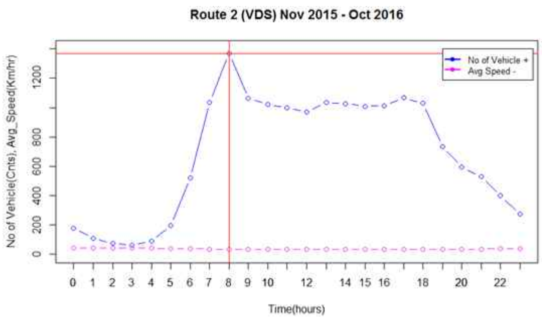

Figure 10. The result of VDS (Route 2) 1-Year Traffic Data.

The blue dotted-dashed line refers to the average of a number of vehicles and it is increased from morning 6 o'clock to evening 6 o'clock. The pink dotted-dashed line refers to the average of average speed of the vehicle and it is decreased at evening 6 o'clock. The maximum value of counts (number of the vehicle) and a minimum value of the average speed of the vehicles are at evening 6 o'clock. So, it is satisfied Method 1 of peak hour identification. Therefore, 18th hour (evening 6 o'clock) is a perfect peak hour on Route 1 of VDS traffic data. Figure 8 clearly shows up the busy hours are from morning 8 o'clock to evening 6 o'clock.

Figure 9 shows the result of RSE (Route 2) 1-Year traffic data. The blue dotted line increases from morning 8 o'clock to evening 6 o'clock, green dotted line increases from morning 7 o'clock to evening 6 o'clock and the pink line decreases at evening 6 o'clock. So, this result is satisfying the Method 1 of peak hour identification. Therefore, the busy hours are from morning 8 o'clock to evening 6 o'clock and the peak hour is at evening 6 o'clock in RSE (Route 2) Traffic Data.

Figure 10 shows the result of VDS (Route 2) traffic data. The blue dotted line refers to the number of vehicles and it's decreased from morning 8 o'clock to evening 6 o'clock and pink dotted line refers to the average speed of the vehicle and its decrease at morning 8 o'clock. So, this result satisfied the Method 1 of peak hour identification. Therefore, in that the busy hours are from morning 8 o'clock to 6 o'clock and the peak hour is at morning 8 o'clock.

5. Conclusions

This study proposed the analysis and peak hour identification of traffic congestion with the help of Kalman Filter and Ensemble Model. The main intent of this study is to prevent roadway traffic congestion with the help of peak hour identification. In traffic data there are two types, one for Roadside Equipment and another for Video Detection System. Both data were collected from RSE device and VDS device. In this traffic data, the Kalman Filter is used to remove the noisy data and for the purpose of implementing the Ensemble model is averaging the data. There are four types of methods to find the peak hour identification. Based on the result, we identified the busiest hours are from morning 8 o'clock to evening 6 o'clock on Route 1 and Route 2. In Route 1, the peak hour of RSE and VDS is at evening 6 o'clock, and in Route 2 the peak hour of RSE is at evening 6 o'clock and VDS is at morning 8 o'clock. Because of, sometimes the RSE and VDS placed in different places on same Route 2. This is the reason to get this type of results in Route 2. The author suggests that this study is really helpful in Suncheon City Corporation.

Generally, daytime is working time for all persons and so, the number of vehicles is increased in the daytime in traffic data. In the result of RSE and VDS data, the peak hour is at evening 6 o'clock. For the reason of, during the daytime, everyone went to their job with different timing (at morning 8 o'clock or 9 o'clock, etc.) and then, naturally many of the people returning their home at evening 6 o'clock (in Korea, evening 6 o'clock is finishing time off work). So, this is the reason for peak hour at evening 6 o'clock on Route 1. Future research focuses on analysis and forecasting the roadway traffic data.

References

[1] R. Arnott, and K. Small, *The economics of traffic congestion*, The Scientific Research Society, Vol. 82, No.5, pp.446-455, Oct. 1994.

[2] R. Arnott, *Does providing information to*

drivers reduce traffic congestion?, Transportation Research Part A: General, Vol. 25, issue 5, pp. 309-318, Sep. 1991.

[3] Y. Sugiyama, M. Fukui, M. Kikuchi, K. Hasebe, A. Nakayama, K. Nishinari, S-I. Tadaki, and S. Yukawa, *Traffic jams without bottlenecks-experimental evidence for the physical mechanism of the formation of a jam*, Vol. 10, Mar. 2008.

[4] S. Rosenbloom, *Peak-period traffic congestion: A state-of-the-art analysis and evaluation of effective solutions*, Transportation, Vol. 7, issue 2, pp. 167-191, Jun. 1978.

[5] C. Sommer, R. German, and F. Dressler, *Bidirectionally coupled network and road traffic simulation for improved IVC analysis*, IEEE, Vol. 10, No. 1, pp. 3-15, Jan. 2011.

[6] V. Shankar, F. Mannering, and W. Barfield, *Effect of roadway geometrics and environmental factors on rural freeway accident frequencies*, Accident Analysis & Prevention, Vol. 27, No. 3, pp. 371-389, Jun. 1995.

[7] Y. Wang, and M. Papageorgiou, *Real-time freeway traffic state estimation based on extended Kalman filter: a general approach*, Transportation Research Part B: Methodological, Vol. 39, No. 2, pp. 141-167, Feb. 2005.

[8] L. Fei, H. B. Zhu, and X. L. Han, *Analysis of traffic congestion induced by the work zone*, Physica A: Statistical Mechanics and its Applications, Vol. 450, pp. 497-505, May 2016.

[9] C. Bian, C. Yuan, W. Kuang, and D. Wu, *Evaluation, classification, and influential factors analysis of traffic congestion in Chinese cities using the online map data*, Vol. 2016, pp. 10, Aug. 2016.

[10] E. D'Andrea, and F. Marcelloni, *Detection of traffic congestion and incidents from GPS trace analysis*, Expert Systems with Applications, Vol. 73, pp. 43-56, May 2017.

[11] K. Shi, B. Di, K. Zhang, C. Feng, and L. Svirchev, *Detrended cross-correlation analysis of urban traffic congestion and NO_2 concentrations inchengdu*, Transportation Research PartD: Transport and Environment, Vol. (inpress), pp. (inpress), Jan. 2017.

[12] L. Wei, D. Hong-ying, *Real-time road congestion detection based on image texture analysis*, Procedia Engineering, Vol. 137, pp. 196-201, 2016.

[13] M. A. Quddus, C. Wang, and S. G. Ison, *Road traffic congestion and crash severity: econometric analysis using ordered response models*, Journal of Transportation Engineering, Vol. 136, No. 5, pp. 424-435, 2010.

[14] C. Wang, M. A. Quddus, and S. G. Ison, *Impact of traffic congestion on road accidents: Aspatial analysis of the M25 motorway in England*, Accident Analysis & Prevention, Vol. 41, pp. 798-808, 2009.

[15] Z. Zheng, S. Ahn, and C. M. Monsere, *Impact of traffic oscillations on freeway crash occurrences*, Accident Analysis & Prevention, Vol. 42, pp. 626-636, 2010.

[16] Kalman filter from Wikipedia, https://en.wikipedia.org/wiki/Kalman_filter, 25 May 2017.

[17] D. Mfinanga, and E. Fungo, *Impact of incidents on traffic congestion in Dar es Salaam city*, International Journal of Transportation Science and Technology, Vol. 2, pp. 95-108, 2013.

[18] S. Roshandel, Z. Zheng, and S. Washington, *Impact of real-time traffic characteristics on freeway crash occurrence: Systematic review and meta-analysis*, Accident Analysis & Prevention, Vol. 79, pp. 198-211, 2015.

[19] Ensemble Modelling from TechTarget, http://searchbusinessanalytics.techtarget.com/definition/Ensemble-modeling, 25 May 2017.

[20] P. S. P. Cowpertwait, and A. V. Metcalfe, *Introductory time series with R*, New York: Springer Dordrecht Heidelberg, 2009.

IoT 환경에서 교통 혼잡에 대한 첨두시 규명

Vasanth Ragu, Saraswathi Sivamani, 이명배, 조현욱, 조용윤, 박장우, 신창선
순천대학교 정보통신공학과

요 약

본 연구에서는 앙상블 모델과 칼만필터를 이용하여 교통혼잡에 대한 첨두시를 분석하고 규명한다. 혼잡제어는 도로교통, 항로교통, 네트워크 등에서 요구되며 본 연구는 도로교통에서의 혼잡제어에 초점을 맞췄다. 첨두시의 규명은 도로에서 교통혼잡을 회피하고 방지하기 위해 필수적인 부분이다. 본 연구에 활용된 도로교통데이터 수집장치는 노변장치(Roadside Equipment, RSE)와 영상검지시스템(Video Detection System, VDS)로 구분되며, 각 데이터는 실제 도로변과 도로요금소 등에 설치된 상기의 장치들로부터 수집된다. 이렇게 수집된 교통데이터에는 오류값이 포함될 수 있다. 수집된 데이터에 존재하는 오류 및 부정확한 정보를 정제하기 위해 칼만필터를 사용했으며, 교통혼잡 규명을 위한 평균데이터를 산출하기 위해 앙상블모델을 적용했다. 그리고 교통데이터에서 정확한 첨두시 규명을 위해 차량수, 시간, 속도에 따른 4개의 다른 모델을 만들어 분석했다. 분석결과를 통해 제시하는 방법을 통해 교통혼잡에 대한 첨두시를 명확히 규명함을 확인했다.

Acknowledgments

This paper was supported by (in part) Sunchon National University Research Fund in 2016.

Vasanth Ragu completed Master of Computer Application (MCA) degree from India. He is pursuing a Doctorate degree in information and Communication Engineering at Sunchon National University in South Korea. His area of interest includes Forecast Model, Ubiquitous Computing, and Big Data Processing.

E-mail address: vasanth4224@scnu.ac.kr

Saraswathi Sivamani completed Bachelor degree in Information Technology from India. She received Master degree on Information and Communication Engineering in South Korea and currently pursuing Doctorate degree in the same department. Her area of interest includes Ontology model, Ubiquitous Computing and Big data processing. She has worked as Senior Engineer, Java Developer for three years in the Tata Consultancy Services, India.
E-mail address: saraswathi86@scnu.ac.kr

Myeongbae Lee completed Bachelor degree in Computer Engineering from Korea. He received Master degree on Computer Science in South Korea. And currently pursuing Doctorate degree in the Information and Communication Engineering. He area of interest includes Advanced Agriculture Technology, IT Convergence, Cloud and Ubiquitous Computing.

E-mail address: lmb@scnu.ac.kr

Hyunwook Cho received Bachelor degree on Computer Science at Korea National Open University in the year 2007. Currently, his Pursuing Doctorate degree on Information and Communication Engineering at Sunchon National University. His area of interest includes Ubiquitous Computing, Wireless sensor, and Big Data Processing.

E-mail address: chohyunwook@scnu.ac.kr

Yongyun Cho received the Ph.D. degree in computer engineering at Soongsil University. Currently, he is an assistant professor of the Department of Information & communication engineering in Sunchon National University. His main research interests include System Software, Embedded Software and Ubiquitous Computing.

E-mail address: yycho@scnu.ac.kr

Jangwoo Park received the B.S., M.S. and Ph.D. degrees in Electronic engineering from Hanyang University, Seoul, Korea in 1987, 1989 and 1993, respectively. In 1995, he joined the faculty member of the Sunchon National University, where he is currently a professor in the Department of Information & Communication engineering. His research focuses on Localization and SoC and system designs and RFID/USN technologies.

E-mail address: jwpark@scnu.ac.kr

Changsun Shin received the Ph.D. degree in Computer Engineering at Wonkwang University. Currently, he is a professor of the Dept. of Information & Communication Engineering in Sunchon National University. His researching interests include Distributed Computing, IoT, Machine Learning, and Agriculture/ICT Convergence.

E-mail address: csshin@scnu.ac.kr