

Saturation Prediction for Crowdsensing Based Smart Parking System

Mihui Kim* and Junhyeok Yun*

Abstract

Crowdsensing technologies can improve the efficiency of smart parking system in comparison with present sensor based smart parking system because of low install price and no restriction caused by sensor installation. A lot of sensing data is necessary to predict parking lot saturation in real-time. However in real world, it is hard to reach the required number of sensing data. In this paper, we model a saturation prediction combining a time-based prediction model and a sensing data-based prediction model. The time-based model predicts saturation in aspects of parking lot location and time. The sensing data-based model predicts the degree of saturation of the parking lot with high accuracy based on the degree of saturation predicted from the first model, the saturation information in the sensing data, and the number of parking spaces in the sensing data. We perform prediction model learning with real sensing data gathered from a specific parking lot. We also evaluate the performance of the predictive model and show its efficiency and feasibility.

Keywords

Crowdsensing, Regression Model, Saturation Prediction, Smart Parking System

1. Introduction

Wireless sensor networks (WSNs) technologies have been attracting a lot of attention from researchers because of the potential use in various fields, i.e., military, environment, city, and so on. However, the realization of WSNs is slow with many initial cost of installing the sensors. Thus, a new service paradigm, crowdsensing, has recently emerged to reduce the initial cost of deploying WSNs and provide the service even when it is difficult or impossible to install the sensors realistically. Crowdsensing is a system that collects sensing information from the public with sensors such as smart phones and smart watches, and provides services to users who requires the sensing information [1,2]. Especially, crowdsensing service can be easily operated in a crowded city where many people with sensors have, that is essential for implementing urban-based Internet services [3,4].

As an application of urban-based Internet services, smart parking system provides saturation information as well as parking lot information, i.e., location, the number of accommodated vehicles, and parking fee. In most existing smart parking systems, saturation information is provided through sensors installed at each parking space in entire parking lot [5]. This method can provide accurate saturation

* This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Manuscript received October 1, 2018; first revision December 5, 2018; accepted April 9, 2019; onlinefirst July 15, 2019

Corresponding Author: Mihui Kim (mhkim@hknu.ac.kr)

* Dept. of Computer Science and Engineering, Computer System Institute, Hankyong National University, Anseong, Korea
(mhkim, junhyeok2723)@hknu.ac.kr

information for the entire parking lot. However, it cannot be used in parking lots without sensor installed, the initial cost for sensor installation is high, and thus smart parking system has low penetration into small parking lots. Using crowdsensing, parking lot saturation information can be provided based on the participation of the parking lot user without installing the sensor. As our previous work [6], we designed a rough architecture of smart parking system based on crowdsensing to replace existing parking systems and only focused on removing privacy information in sensing data.

In the system of [6], it is assumed that the sensing data provider provides the service provider server with saturation for 4 to 6 parking spaces on average through the photograph information. Accordingly, in order for the service provider server to provide the actual saturation information of the entire parking lot to the user, a plurality of sensing data is required. The amount of required sensing data increases in proportion to the size of the parking lot (the number of parking spaces). Moreover, since the parking lot saturation rapidly changes, the amount of sensing data required to provide the actual saturation information of the entire parking lot in real time further increases. The increase in the amount of sensing data required by the server to the user means an increase in the compensation budget paid to the sensing data provider. Therefore, in reality, it is difficult to satisfy the amount of sensing data required to calculate the actual saturation of the entire parking lot. Thus, in order to operate the smart parking system based on crowdsensing in reality, it is ideal to provide whole saturation prediction information based on the sensing data including the saturation information of a part of the parking lot.

In this paper, we propose a method to process sensing data provided by a sensing information provider and convert it into saturation information. In particular, we propose a prediction model that predicts the total saturation based on the saturation information of a part of the parking lot provided by the user using regression analysis. That is, we design the information processing module of service provider system in detail, based on a rough architecture of smart parking system based on crowdsensing, our previous work. We perform prediction model learning with real sensing data gathered from a specific parking lot, and evaluate the performance of the prediction model to show its efficiency and feasibility of the system.

This paper is organized as follows. Section 2 presents related work. In Section 3, we propose a prediction model for smart parking system. Section 4 shows the results of prediction model learning. Section 5 evaluates our proposed model, and Section 6 concludes this paper.

2. Related Work

2.1 Crowdsensing

Fig. 1 shows a flow from information request to information provision in crowdsensing service. The steps are as follows: (1) 'User A' who needs information sends a request for information to the server. (2) The server sends a request for providing the sensing data to the 'user B', which can provide the information required by the 'user A'. (3) 'User B' senses the data and (4) sends it to the server. (5) The server processes the sensing data provided by the 'user B' and makes the information necessary for the 'user A'. (6) The server transmits the processed information to 'user A'.

If it is difficult or impossible to install a sensor on a system that needs to use sensor information, crowdsensing can be used to implement the system without installing a sensor. There are two types of crowdsensing: participatory crowdsensing and persistent crowdsensing. The participatory crowdsensing

sends the request to the user when an information needs, and the user who receives the request decides whether to participate [7]. At this time, the user providing the sensing data obtains compensation for the sensing data from the server [8]. In the persistent crowdsensing, a server continuously collects information from users [9].

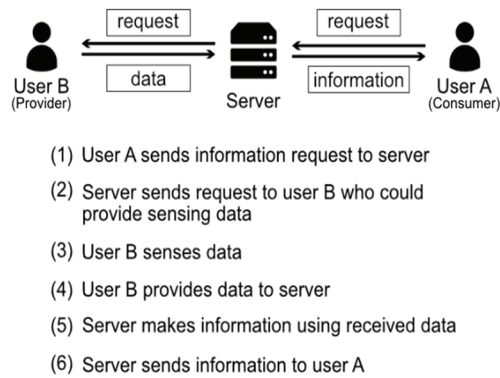


Fig. 1. Crowdsensing service flow.

In the smart parking system proposed in this paper, a participatory crowdsensing is applied to send sensing data provision request to the user in the place where the sensing data is needed, and the parking lot saturation information is predicted and provided based on the sensing data provided by the user [6].

2.2 Research Trends of Crowdsensing

Crowdsensing technologies have attracted much attention as an alternative to WSNs that has a large initial facility cost, and related research is proceeding actively. We introduce research trend of crowdsensing technologies. There are incentive mechanism research for realizing crowdsensing, security problems to assure data reliability or protect personal information concerned in sensing data sharing, and related application research.

2.2.1 Incentive mechanism

As an open issue of crowdsensing, incentive mechanism is an important part to incentivize users to be workers providing sensing services. That is, in real world, users might not provide sensing data if crowdsensing system would not pay adequate incentive. Moreover, if every user who provide sensing data gets the same amount of incentive, users would not try to provide high-quality sensing data. Thus researchers in [10] proposed an incentive mechanism based on a quality-driven auction, and included a probabilistic model to evaluate the reliability of the submitted data through Wi-Fi fingerprint-based indoor localization. Meanwhile, if users could not get satisfactory incentive, users would not provide data. The proposed incentive mechanism QUOIN in [11] for application requirements and guarantees each participant achieves a satisfactory level of profits. Similarly adequate incentive in [12,13] is decided by estimating data quality and participants are paid incentive differently depending on estimation results. Effort to collect sensing data and the trustiness of users in [14] are suggested as a way to decide the amount of incentive. On the contrary, some researches consider a way to operate unpaid crowdsensing system [15].

2.2.2 Security

There are two main security issues in crowdsensing: assuring data reliability and preserving the privacy of sensing data providers. Firstly, crowdsensing system is operated with sensing data provided by participants to make meaningful information, and thus the success of the crowdsensing service depends on the reliability of the data. Moreover, if there are some users have malicious purpose, i.e., malfunction of crowdsensing system or ill-gotten incentive, crowdsensing system could not provide meaningful information using sensing data. Researchers in [16] proposed a framework FIDC to defend collusion attacks that malicious participants collaboratively send fake information to mislead the system and to improve data credibility. In [17], researchers focused on detecting these Sybil identities who falsify multiple identities and negatively influence the effectiveness of sensing data, through trust management system. That is, many researchers have concentrated on designing methods to block invalid data provision and assure data reliability through data quality estimation and user trustiness management.

Secondly, many sensing data in crowdsensing system include the private information of data providers, e.g., location [18], provider information [19], etc. In crowdsensing system without proper privacy preservation mechanisms, a user would not participate in sensing despite pronounced incentive. Thus, many researchers [18-21] concentrate on taking countermeasure for privacy issues to make users participate in providing sensing data. The mechanism proposed in [20] utilizes a random-sampling based on the privacy auction for data in order to guarantee the data privacy of participants.

2.2.3 Application

Crowdsensing system could be applied to various sensing data-based system, especially for the cases which is hard to install sensors and build sensor network, i.e., smart city [22], wireless LAN monitoring [23,24], data recognition [25], etc. vCity Map [22] visualizes the city environment for two kinds of information: sound and road conditions. Pazl [23] provides indoor WiFi monitoring system with

Table 1. Summarization of main researches

Subjects	Ref.	Feature	Method
Incentive mechanisms	[10]	Quality-driven auction	Wi-Fi fingerprint-based indoor localization for data reliability evaluation
	[11]	Data quality and usability estimation	Stackelberg game model for guaranteeing the satisfaction of each participant though
Security: assuring data reliability	[16]	Defend collusion attacks	Spatial correlation of sensing data and the correlation between sensing data and provenance knowledge
	[17]	Detect Sybil identities	Trust management with the knowledge of traffic volume, signal strength, and network topology
Security: privacy preservation	[18]	Location privacy	Differential geo-obfuscation
	[19]	Protect provider information (i.e., location)	Cloud-based agents with decentralized quadtree
	[20]	Data privacy of participants	Random-sampling based on the privacy auction
Application	[22]	Smart city map	Visualization of sound and road conditions
	[23,24]	Wi-Fi monitoring	Individual measurements indoor [23] and urban [24]
	[25]	Data recognition	Cycles of crowd-querying and feedback
	Our work	Smart parking system providing saturation information	Combined prediction model

individual measurements taken from participant phones. Similarly researchers [24] presented crowdsensing based urban WiFi characterization, i.e., the presence of deployed WiFi AP, used channels, and location. The authors [25] designed an algorithm for finding regions of interest in mobile crowdsensing, through utilizing cycles of crowd-querying and feedback.

In this paper, we design a crowdsensing based smart parking system, and provide the saturation prediction model of parking lot from the image information of participants. Table 1 summarizes the main researches.

2.3 Regression Analysis

Regression analysis is an analytical method that evaluates the performance of a model that represents the relationship between two or more variables [26]. The variables to be predicted through regression analysis are called dependent variables and the factor variables that affect dependent variables are called independent variables. In simple regression analysis, one independent variable is used to predict the value of a dependent variable. In multiple regression analysis, two or more independent variables are used for the value of a dependent variable. If the relationship between the independent variable and the dependent variable forms a nonlinear curve, the dependent variable can be predicted using a polynomial model with the form of Eq. (1), where y denotes a dependent variable, x is an independent, w_i s are coefficients ($0 \leq i \leq d$), and d is the degree of polynomial. The degree d is decided depending on the relationship between independent and dependent variables.

$$y = w_0 + w_1x + w_2x^2 + w_3x^3 + \dots + w_dx^d \quad (1)$$

In this paper, we make a time-based polynomial regression model that predicts the total saturation of parking lot with time as independent variables. For example, the model predicts the saturation of *ParkingLotA* at 20:00 as 0.8 (saturation 1 means full). We then create a sensing data-based linear multiple regression model such as Eq. (2) with saturation value x_1 predicted by time-based regression model, the saturation degree x_2 of parking lot part included in sensing data, and the number of parking spaces x_3 in the sensing data. The sensing data-based linear multiple regression model predicts the degree of saturation y of the parking lot based on the user provided sensing data. w_i s ($0 \leq i \leq 3$) are a weight that indicates the effect of each independent variable on the dependent variable

$$y = w_0 + w_1x_1 + w_2x_2 + w_3x_3 \quad (2)$$

Using the generated regression model, we predict the saturation of the parking lot based on the sensing data provided by the user (Provider in Fig. 1) and provide the saturation information to the user (Consumer in Fig. 1). The performance of the generated regression model is evaluated by the mean squared error (MSE) value [27], which represents the error between the predicted value and the actual saturation of the regression model, and the R-squared (R^2) score, which represents the degree to which the predicted value of the regression model is suitable to represent the actual saturation [28].

2.4 Previous Work [6]

Fig. 2 is a system structure proposed in [6], as our previous work, to provide crowdsensing based smart

parking system. The system consists of two entities: users who provide sensing data or use information, and a service provider that provides and operates a service. The user application includes an Information Sensing module, an Information Requesting/Using module for requesting information provided to the service provider or for using the provided information, an Information Processing module for removing personal information from the sensing data, and an Incentive Managing module that manages compensation for sensing data. The server of service provider is composed of a User Managing module, an Information Processing module for detecting an empty parking space and saturation prediction, and an Incentive Managing module for managing the compensation paid to the sensing data provider.

The processing flow between two entities is shown in Fig. 3. The service user application periodically sends location information to the server. (1) When the server of service provider receives a request for parking information from a user A, (2) it searches for users in the corresponding area, and (3) requests information provision. (4) (5) User B requested to provide information removes personal information (i.e., image obfuscation process) such as a license plate before transmitting the photographed parking lot image to the server. (6) The server calculates the position of the empty parking space and the parking lot saturation information through the received information, and (7) provides the information to the user A.

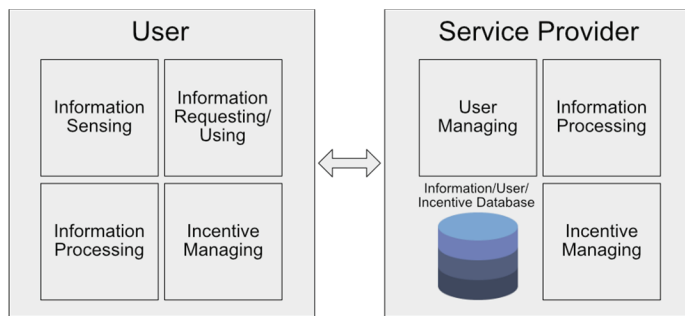


Fig. 2. System structure [6].

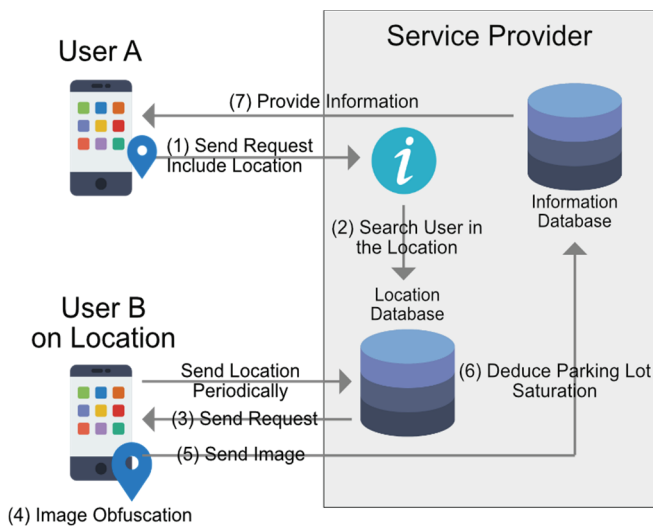


Fig. 3. System flow [6].

Based on the previous work [6], in this paper, we propose a prediction model that the Information Processing module of the service provider server uses to predict the total saturation based on sensing data supplied from the users. We also evaluate the performance of the prediction model and show its feasibility.

3. Information Processing Module for Prediction

In this paper, we focus on designing an Information Processing module in service provider server on smart parking system of our previous work [6]. Fig. 4 shows the structure of the information processing module, including a parking Block Size Decision module for determining the number of parking spaces included in one parking block, a Data Processing module for converting the sensing data into data for predicting saturation, and a Saturation Prediction module for predicting the total saturation based on the data processed in the data processing module.

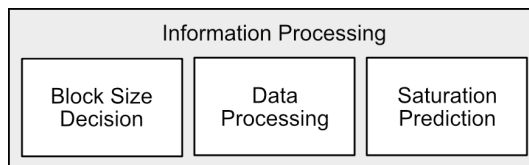


Fig. 4. Information Processing module structure.

3.1 Block Size Decision

By separating the parking lots into several blocks with a certain number of parking spaces, it is possible to specify more precisely which block the empty parking space is in. Moreover, by separating blocks, different saturation patterns can be predicted and users can be guided to approach parking blocks with low saturation.

The smaller the number of parking spaces included in one parking block, the more accurate the position of the vacant parking space can be provided to the user. However, if the number of parking spaces included in one parking block is too small, the amount of sensing data needed to predict the total saturation increases, losing the benefit of using the prediction model. We will evaluate the prediction performance by changing the number of parking spaces included in one parking block, as future work.

3.2. Data Processing

The user provides the server with a partial image of the parking lot. The server detects the block number and the parking status in the provided image. The cascade detection algorithm used in [6] is used for the detection of block number and parking status.

timestamp	block	length	state	saturation
-----------	-------	--------	-------	------------

Fig. 5. Sensing data structure.

The sensing data structure detected from the image provided by the user is as shown in Fig. 5. The value of the *timestamp* is the time at which the information was provided. The *block* indicates to which block the parking space indicated in the image belongs. The value of the *length* is the number of parking spaces included in the image. The value of the *state* is a bit string indicating the parking status of the parking space included in the image. Sensing data providers may be asked to transmit information that may be a reference point, such as a number on a parking pillar or on a wall. Thus, the state bits can consist with the reference position information. An empty parking space is expressed as the state value 0, and a parking space in which the vehicle is parked is represented by 1. The value of the *saturation* is the degree of saturation of the portion of the parking lot shown in the image.

$$S = \frac{\sum_{n=1}^l a_n}{l} \quad (3)$$

Eq. (3) represents the method of calculating saturation, which is the value of the saturation S . l is the number of parking spaces included in the image and is equal to the value of the length field in the sensing data structure. a_n is the n^{th} matrix representing the parking state of the parking spaces included in the image, and is equal to the value of the state field in the sensing data structure.

If the sensing data provided by several users is data for the same part of the parking lot, the accuracy of the prediction result may be degraded by predicting saturation based on the actual parking state and other data. Therefore, redundancy is eliminated by using the state field value of the sensing data structure. Fig. 6 is the example diagrams illustrating a method of eliminating redundancy for sensing data A and sensing data B provided by different users. These bits strings for sensing data consist with reference position information (e.g., parking lot number A0, A1, and so on). In Fig. 6(a), when the state bit string of the sensing data A is included in the state bit string of the sensing data B, the sensing data A is deleted. In Fig. 6(b), if there is an overlap between the state bit string of the sensing data A and the state bit string of the sensing data B, the overlapping portion is deleted from the state bit string of the sensing data B and the two are collected. Based on this, the data processing module eliminates redundancy and keeps the data up-to-date if changes occur in a short time.

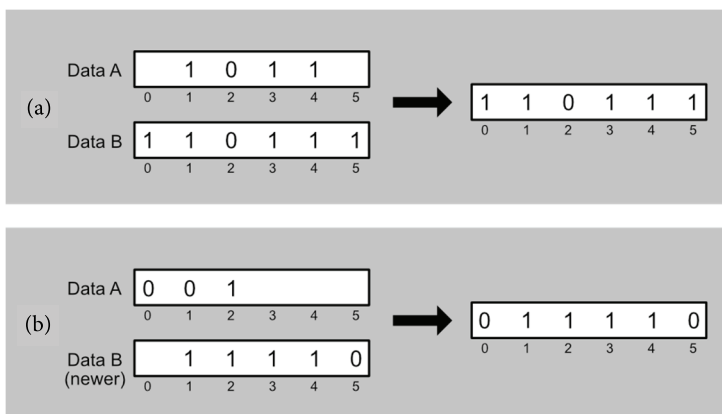


Fig. 6. Examples of removing duplicate sensing data. (a) Case that one string is included in another string, (b) Case that an existing string overlaps a new string.

3.3 Saturation Prediction

According to the location of the parking lot, the shape of the hourly saturation is different. For example, a parking lot in a residential area shows high saturation before and after work hours on weekdays, and the parking lot of a large mart shows high saturation on a weekend afternoon. Therefore, the prediction model uses the location and time of the parking lot as variables. The location of the parking lot is set based on the location specified in the data supply request. The time is the value of the timestamp field included in the sensing data. Since the prediction model using the location and time of the parking lot depends on the existing data, the accurate saturation of real time cannot be predicted. Therefore, in order to obtain accurate prediction results based on real-time data, the saturation degree of the parking lot part included in the sensing data is used as a variable. At this time, the saturation degree of the parking lot portion is set to the saturation field value of the sensing data. The server provides the user with a block number and saturation prediction for the parking block with an empty parking space. The real-time accuracy of the information can be evaluated and provided according to how much information provided at this time refers to the real-time sensing data.

4. Prediction Model Learning

The prediction model of this system consists of a polynomial regression model that predicts the degree of saturation with time as a variable and a linear regression model that predicts the degree of saturation of the part of the parking lot that is not included in the sensing data with the information of the parking lot part included in the sensing data. A time-based prediction model predicts the degree of saturation observed at a specific time in a specific place, and then increases its accuracy with a sensing data-based prediction model. The learning data for the prediction model learning was collected hourly for 2 days at the same place with the hourly saturation degree for 52 parking spaces in total. We use TensorFlow [29] to generate prediction models by artificial neural network based on sensing data. In artificial neural network, we set up the saturation degree of the part included in the sensing data, the size of the sensing data, and the time-based predicted saturation as the input variables. We also set up the saturation degree of the part not included in the sensing data as the dependent variable. Using the Gradient Descent algorithm, we learn until the error value dropped below $1e-5$. We also use scikit-learn [30] to calculate the accuracy of predicted values, R^2 and MSE values through the generated prediction model.

Fig. 7 is a graph showing the real saturation information of the hourly hour angle at the same place and the time-based polynomial regression model learned based on the actual saturation information. As a result of the prediction of the degree of saturation according to time, the accuracy is relatively high in the time zone where the daily saturation change is not large. However, it can be confirmed that the accuracy is lowered when the saturation aspect appears slightly different every day. In this case, a linear regression model based on the user-provided sensing data can be used together to obtain a higher accuracy saturation prediction value.

The degree of the polynomial used in the time-based polynomial regression model is determined by the MSE value of the learned regression model and the score of R^2 . The MSE value and R^2 score according to the order of the polynomial are shown in Fig. 8. When the degree is above 5, the lowest MSE value and the highest R^2 score are shown. The higher the degree of the polynomial, the more time it takes to predict

the degree of saturation. Therefore, it is advantageous to select a lower order if the change in the MSE and R^2 values is small. The degree of the polynomial is set to 5 in the time-based polynomial regression model proposed in this paper.

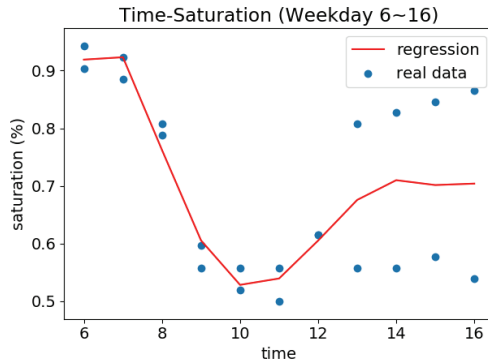


Fig. 7. Time-saturation graph.

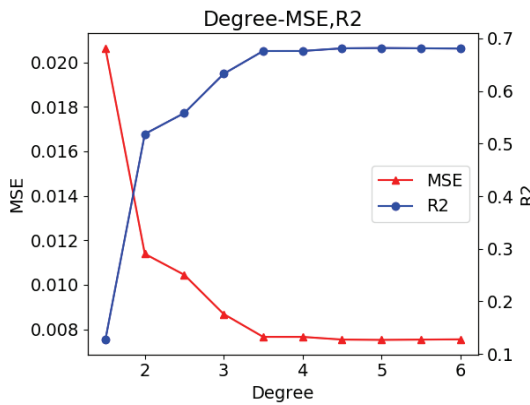


Fig. 8. Degrees of MSE and R^2 .

The sensing data-based prediction model predicts the degree of saturation of the parking space not included in the sensing data, based the predicted saturation degree predicted by the time-based prediction model, the degree of saturation of a part of the parking lot included in the sensing data, and the number of parking spaces included in the sensing data as variables.

Fig. 9 shows the MSE value according to the size of the sensing data (the number of parking spaces included in the sensing data) input to the sensing data-based prediction model. The results show that the sensing data generation, learning, and MSE calculation of the predicted values are repeated 10 times and averaged. When the size of the sensing data is 0, it means that the saturation degree is predicted using only the time-based prediction model. The MSE value when the sensing data-based prediction model is used together is lower than the MSE value when only the time based prediction model is used. Moreover, it can be confirmed that the MSE value decreases as the number of parking spaces included in the sensing data increases and the sensing data, and the sensing data supplements the lack of only time-based prediction model. This means that the prediction accuracy is higher when the sensing data based prediction model is used together than when using only the time based prediction model.

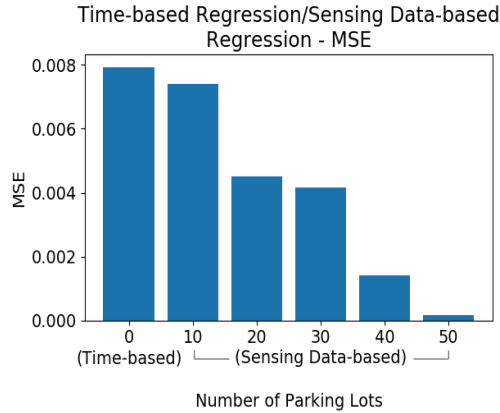


Fig. 9. MSE of time-based and sensing data-based models.

5. Performance Evaluation

In this section, we evaluate the performance of our system model depending on the block size and whether or not the outlier is removed.

5.1 Parking Block Size

Fig. 10 is a graph showing the MSE value according to the number of parking spaces included in one parking zone and the required minimum data quantity. The total number of parking spaces in the whole parking lot is 52, which is the result of separation into 1 parking block, 2 parking blocks and 4 parking blocks, respectively. The minimum amount of sensing data required to predict the overall parking saturation increases in proportion to the number of parking spaces. The MSE value of the saturation prediction result is lowest in the model where the entire parking lot is divided into two parking blocks.

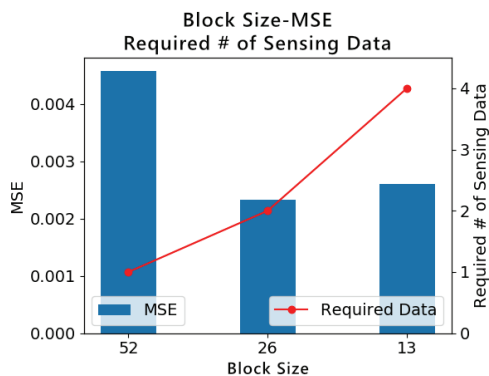


Fig. 10. Block size-regression performance.

If the parking block is not divided (52 parking spaces per parking block), the MSE value of the predicted model is relatively high, while the total saturation can be predicted with only one sensing data. If the entire parking lot is divided into 4 parking blocks (13 parking spaces per parking block), the MSE value

of the predicted model is relatively low, but the minimum amount of sensing data required to predict the total saturation is high. The lowest MSE value was obtained when the entire parking lot was divided into 2 parking blocks (26 parking spaces per parking block). In the case with 2 parking blocks, the minimum number of sensing data required to predict the total parking lot saturation was also smaller than the case that is divided into 4 parking spaces. If the user participation rate is low, the parking zone size with the lowest MSE value of the prediction model can be selected within the range of satisfying the required data quantity.

It can be expected that the MSE value of the prediction model decreases as the parking blocks are separated. In addition, when the parking lots are divided into a plurality of blocks, it is possible to more precisely specify the position of the vacant parking space and provide the information to the user. Therefore, it is more efficient to separate parking lots into multiple parking blocks if the user has sufficient data participation rate.

5.2 Outlier Removal

In this subsection, we compare the results with and without outlier removal. Results are taken with 2,000 sensing data in where the block size is 52. In case of outlier removal, we remove the top 5% of the difference between the saturation and actual saturation in the generated sensing data. Fig. 11 is a graph showing the MSE value of the regression model that is learned based on the learning data with and without the anomaly values. The results are averaged over repeated 10 times. The outlier removal is performed by deleting the upper 5% of the data based on the difference between the degree of total saturation and the degree of partial saturation included in the training data. As a result, the MSE value when the outlier values are removed is smaller than the value when they are not.

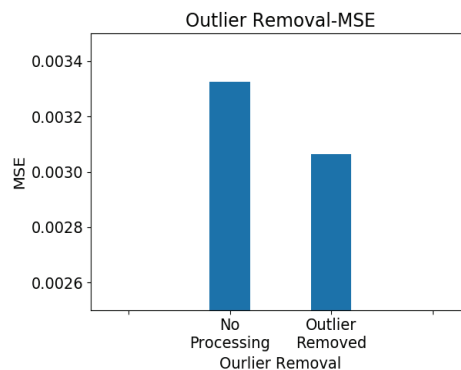


Fig. 11. Outlier removal-MSE.

If the size of the parking block is small, the probability of a large difference between the degree of total saturation and the degree of partial saturation is relatively low. Therefore, higher prediction accuracy can be expected by using the prediction model learned based on the learning data from which the outlier values are removed. However, if the size of the parking block is large, the difference between the degree of total saturation and the degree of partial saturation is relatively high. Therefore, prediction accuracy can be lowered by using prediction models learned based on learning data from which outliers have been removed. In this case, it is safe not to apply the manipulation to the learning data because the difference

between the MSE values of the prediction model in which the outliers are removed and the prediction model in which they are not is not large.

6. Conclusion

In this paper, we propose a method of predicting total saturation based on sensing data in information processing module of crowdsensing based smart parking system. We have implemented a predictive model and performed the prediction model learning. The comparison of MSE between time-based prediction model and combined prediction model showed that sensing data-based prediction model makes the high prediction accuracy, because the sensing data provides insufficient information in time-based prediction model. As future research, we will design the compensation mechanism of the proposed system and develop a realistic crowdsensing based smart parking system. For example, even if users do not consciously participate in sensing data provisioning, the application can use devices that can collect sensing data, e.g., the black box or rear camera of vehicle, CCTV, etc. By setting the vehicle cameras to periodically send photos when the car is parked in public parking lots, users will be able to receive incentives for crowdsensing.

Acknowledgement

This research was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No.2018R1A2B6009620).

References

- [1] J. Burke, D. Estrin, M. Hansen, A. Parker, N. Ramanathan, S. Reddy, and M. B. Srivastava, "Participatory sensing," Center for Embedded Networked Sensing (CENS), University of California, Los Angeles, CA, 2006.
- [2] A. Jian, G. Xiaolin, Y. Jianwei, S. Yu, and H. Xin, "Mobile crowd sensing for internet of things: a credible crowdsourcing model in mobile-sense service," in *Proceedings of 2015 IEEE International Conference on Multimedia Big Data*, Beijing, China, 2015, pp. 92-99.
- [3] M. Mishbah, D. I. Sensuse, and H. Noprisson, "Information system implementation in smart cities based on types, region, sub-area," in *Proceedings of 2017 International Conference on Information Technology Systems and Innovation (ICITSI)*, Bandung, Indonesia, 2017, pp. 155-161.
- [4] X. Lu, B. Chen, C. Chen, and J. Wang, "Coupled cyber and physical systems: embracing smart cities with multistream data flow," *IEEE Electrification Magazine*, vol. 6, no. 2, pp. 73-83, 2018.
- [5] S. O. Yoo, H. J. Oh, and K. S. Oh, "Design and implementation of parking information support system for inner parking lot based on microprocessor," *Journal of the Korea Society of Computer and Information*, vol. 15, no. 1, pp. 51-59, 2010.
- [6] J. Yun and M. Kim, "Smart parking system using mobile crowdsensing: focus on removing privacy information," in *Proceedings of the KIPS Spring Conference*, 2018, pp. 32-35.
- [7] H. Xiong, D. Zhang, G. Chen, L. Wang, V. Gauthier, and L. E. Barnes, "iCrowd: near-optimal task allocation for piggyback crowdsensing," *IEEE Transactions on Mobile Computing*, vol. 15, no. 8, pp. 2010-2022, 2016.

- [8] M. Musthag, A. Raij, D. Ganesan, S. Kumar, and S. Shiffman, "Exploring micro-incentive strategies for participant compensation in high-burden studies," in *Proceedings of the 13th International Conference on Ubiquitous Computing*, Beijing, China, 2011, pp. 435-444.
- [9] R. K. Ganti, F. Ye, and H. Lei, "Mobile crowdsensing: current state and future challenges," *IEEE Communications Magazine*, vol. 49, no. 11, pp. 32-39, 2011.
- [10] Y. Wen, J. Shi, Q. Zhang, X. Tian, Z. Huang, H. Yu, Y. Cheng, and X. Shen, "Quality-driven auction-based incentive mechanism for mobile crowd sensing," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 9, pp. 4203-4214, 2015.
- [11] K. Ota, M. Dong, J. Gui, and A. Liu, "QUOIN: Incentive mechanisms for crowd sensing networks," *IEEE Network*, vol. 32, no. 2, pp. 114-119, 2018.
- [12] D. Yang, G. Xue, X. Fang, and J. Tang, "Incentive mechanisms for crowdsensing: crowdsourcing with smartphones," *IEEE/ACM Transactions on Networking (TON)*, vol. 24, no. 3, pp. 1732-1744, 2016.
- [13] R. I. Ogie, "Adopting incentive mechanisms for large-scale participation in mobile crowdsensing: from literature review to a conceptual framework," *Human-centric Computing and Information Sciences*, vol. 6, article no. 24, 2016.
- [14] X. Zhang, Z. Yang, W. Sun, Y. Liu, S. Tang, K. Xing, and X. Mao, "Incentives for mobile crowd sensing: a survey," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 1, pp. 54-67, 2016.
- [15] R. M. Borromeo and M. Toyama, "An investigation of unpaid crowdsourcing," *Human-centric Computing and Information Sciences*, vol. 6, article no. 11, 2016.
- [16] T. Zhou, Z. Cai, K. Wu, Y. Chen, and M. Xu, "FIDC: a framework for improving data credibility in mobile crowdsensing," *Computer Networks*, vol. 120, pp. 157-169, 2017.
- [17] S. H. Chang and Z. R. Chen, "Protecting mobile crowd sensing against Sybil attacks using cloud based trust management system," *Mobile Information Systems*, vol. 2016, article ID. 6506341, 2016.
- [18] L. Wang, D. Yang, X. Han, T. Wang, D. Zhang, and X. Ma, "Location privacy-preserving task allocation for mobile crowdsensing with differential geo-obfuscation," in *Proceedings of the 26th International Conference on World Wide Web*, Perth, Australia, 2017, pp. 627-636.
- [19] I. Krontiris and T. Dimitriou, "Privacy-respecting discovery of data providers in crowd-sensing applications," in *Proceedings of 2013 IEEE International Conference on Distributed Computing in Sensor Systems*, Cambridge, MA, 2013, pp. 249-257.
- [20] M. Zhang, L. Yang, X. Gong, and J. Zhang, "Privacy-preserving crowdsensing: privacy valuation, network effect, and profit maximization," in *Proceedings of 2016 IEEE Global Communications Conference (GLOBECOM)*, Washington, DC, 2016, pp. 1-6.
- [21] I. J. Vergara-Laurens, L. G. Jaimes, and M. A. Labrador, "Privacy-preserving mechanisms for crowdsensing: survey and research challenges," *IEEE Internet of Things Journal*, vol. 4, no. 4, pp. 855-869, 2016.
- [22] Y. Tobe, I. Usami, Y. Kobana, J. Takahashi, G. Lopez, and N. Thepvilojanapong, "vcity map: crowdsensing towards visible cities," in *Proceedings of 2014 IEEE SENSORS*, Valencia, Spain, 2014, pp. 17-20.
- [23] V. Radu, L. Kriara, and M. K. Marina, "Pazl: a mobile crowdsensing based indoor WiFi monitoring system," in *Proceedings of the 9th International Conference on Network and Service Management (CNSM)*, Zurich, Switzerland, 2013, pp. 75-83.
- [24] A. Farshad, M. K. Marina, and F. Garcia, "Urban WiFi characterization via mobile crowdsensing," in *Proceedings of 2014 IEEE Network Operations and Management Symposium (NOMS)*, Krakow, Poland, 2014, pp. 1-9.
- [25] S. W. Loke, "Heuristics for spatial finding using iterative mobile crowdsourcing," *Human-centric Computing and Information Sciences*, vol. 6, article no. 4, 2016.
- [26] E. J. Williams, "Linear hypotheses: regression," in *International Encyclopedia of Statistics*. New York, NY: The Free Press, 1978, pp. 523-541.
- [27] E. L. Lehmann and G. Casella, *Theory of Point Estimation* (2nd ed.). New York, NY: Springer Science & Business Media, 1998.

- [28] R. G. D. Steel and J. H. Torrie, Principles and Procedures of Statistics: With Special Reference to the Biological Sciences. New York, NY: McGraw-Hill, 1960.
- [29] TensorFlow, <https://www.tensorflow.org/>.
- [30] Scikit-learn, <http://scikit-learn.org/stable/>.



Mihui Kim <https://orcid.org/0000-0002-4896-7400>

She received the B.S. and M.S. degrees in Computer Science and Engineering from Ewha Womans University in Korea, in 1997 and 1999, respectively. During 1999–2003, she stayed in Switching & Transmission Technology Lab., Electronics and Telecommunications Research Institute (ETRI) of Korea to develop MPLS System and the 10 Gbps Ethernet System. She also received the Ph.D. degree in Ewha Womans University in 2007. She was a postdoctoral researcher of the department of computer science, North Carolina State University from 2009 to 2010. She is currently an associate professor of the Department of Computer Science and Engineering, Hankyong National University in Korea. Her research interests include security and efficient protocol design in IoT and crowdsensing system.



Junhyeok Yun <https://orcid.org/0000-0001-6240-4455>

He is attending for B.S. in Department of Computer Science and Engineering at Hankyong National University. His current research interests include crowdsensing, blockchain and machine learning.