Artificial Intelligence for Transportation: Advice, Interactivity and Actor Modeling: Papers from the 2015 AAAI Workshop

Pricing Procedure in Accordance with Characteristic of Parking Utilization — Analysis Example of Massive Parking Accounting Data

Yuichi Enoki
Department of
Computer Science and Engineering,
Graduate School of Engineering,
Nagoya Institute of Technology

Ryo Kanamori
Institute of Innovation for
Future Society,
Nagoya University

Takayuki Ito
School of Techno-Business Administration,
Graduate School of Engineering,
Nagoya Institute of Technology

Abstract

In most urban area, traffic issue related to parking (e.g. economic (time) and environmental loss in finding a parking space) has been significant, and parking management strategies to set optimal price are often necessary for the parking agencies. In order to revise parking fee appropriately without a reduction of parking demand, a pricing procedure in accordance with characteristic of parking utilization is expected. On the other hand, a large amount of parking data is accumulated automatically with introduction of online parking systems. In this study, we analyze massive parking accounting data, whose data size is 22.5 million accounting data in the past year about 1,050 parking lots, and discuss the characteristics of parking utilization. Moreover parking duration model is developed from the accounting data for each cluster to estimate the parking demand (i.e., parking time) after changing price. As an example of appropriate price procedure, we evaluate the setting of an upper limitation of parking charge with parking demand patterns calculated from the accounting data and the parking duration model.

1 Introduction

Pay-by-the-hour parking spaces are actively used in downtown areas due to enforcing tighter controls on illegal parking. Considering characteristics of parking utilization is important for appropriate pricings of pay-by-the-hour parking spaces. The appropriate pricings according to users' demands are necessary for increasing users' convenience and profitability of parking lots. Currently, a parking charge is set by investigating utilization characteristics for each parking lot. However, there are problems in that variation of revenue becomes large because of the difference in ability for setting an appropriate parking charge, or the load of investigating all parking lots is large.

The purpose of this paper is to propose a procedure of pricing based on data analysis by pay-by-the-hour parking data. We consider utilization characteristics of each parking lot with parking accounting data. Moreover, we calculate revenue by simulating changes of parking time in accordance with changing pricing based on characteristics of parking utilization. In the estimation of revenue by the simulation, it becomes possible to set parking charges by trial and error. Therefore, the difference in pricing ability can be reduced. In addition, we can grasp utilization characteristics based on data analysis to reduce the load of investigating the parking lots. Big Data for analysis is parking accounting data on the 1,050 parking places around Nagoya City provided by Meitetsu Kyosho Corp. This data set contains 22.5 million accounting data from October 1, 2011 to October 3, 2012.

The structure of this paper is explained below. In section 2, we analyze parking data. In detail, we analyze characteristics of parking utilization by cluster analysis, influence of a charge revision on parking time by survival analysis, and change point of utilization characteristics by cluster analysis. In section 3, we propose a more appropriate pricing from the analysis results in section 2, and consider appropriateness of the pricing by simulation. In section 4, we give an overview of the related work. In section 5, we conclude this paper.

2 Analysis of massive parking data

2.1 Parking data

Pay-by-the-hour parking data provided by Meitetsu Kyosho Corp. is used for analysis in this study. We take in about 11 GB of data provided in the CSV format into a MySQL database Such information as parking location, parking capacity, parking fee, user's payment method (cash, credit card and points, etc.) and parking time can be retrieved from this data. Data for analysis is parking accounting data on the 1,050 parking places around Nagoya City. This data set contains 22.5 million accounting data from October 1, 2011 to October 3, 2012.

2.2 Characteristics of parking utilization

(1) Characteristic variables of parking utilization The parking lots are divided into eight clusters by k-means clustering for classifying the parking into each utilization characteristic. Variables include 1) utilization rate of parking, 2) parking time, 3) availability of commuter pass, 4) utilization rate of point card, 5) discount rate of coupon, and 6) discount rate by maximum charge. The reason for selecting each variable and meaning of variables are explained below.

Utilization rate of parking The utilization rate of parking is used for expressing how actively parking is used. Three variables are used in the utilization rate of parking. A day is divided into four time zones such as 1-7 o'clock, 7-13 o'clock, 13-19 o'clock, and 19-25 o'clock. Then, the average of the utilization rate of parking is calculated for each day of the week. For example, if 50 percent of the parking lot is used between 1 and 4 o'clock, and 100 percent is used between 4 and 7 o'clock, the utilization rate of parking between 1 and 7 o'clock is 75 percent ((3 hours $\times$ 0.5 + 3 hours $\times$ 1.0) $\div$ 6 hours = 75 percent). 28 variables are calculated by dividing time (4 time zones $\times$ 7 days = 28 variables). 28 variables are compressed into 3 variables by principal component analysis (PCA) because 28 variables incur excessive classifying. As a result of PCA, 1) a variable indicating a level of the utilization rate, 2) a variable indicating difference in utilization rate of parking between daytime and nighttime, and 3) a variable indicating difference in utilization rate of parking between weekdays and holidays are derived. These three variables are variables in the utilization rate of parking.

Parking time Parking time variables are used for expressing whether utilization is long-term or not. Three variables are used in the parking time variables. In the parking time variables, a day is divided into four time zones such as 1-7 o'clock, 7-13 o'clock, 13-19 o'clock, and 19-25 o'clock in the same way as the utilization rate of parking. For example, if a user enters a parking lot between 1 and 7 o'clock, then the user has been parking for 10 hours, and the variable indicating time between 1 and 7 o'clock is assigned 600 minutes. All accounting data are handled in this process. Then, the average of the parking time is calculated for each time zone and each day of the week in the same way as the utilization rate of parking. In the parking time variables, 28 variables are calculated in the same way as the utilization rate of parking. These variables are compressed into three variables by PCA. As a result of PCA, 1) a variable indicating length of parking time, 2) a variable indicating difference in parking time between early morning and night, and 3) a variables indicating difference in parking time between daytime and nighttime are derived. These three variables are variables in the parking time.

Availability of commuter pass If a parking lot is available for a commuter pass, the value of the availability of the commuter pass variable is 1, otherwise, the value is 0. The availability of commuter pass variable is a variable indicating characteristics of commuter pass users.

Utilization rate of point card The value of the utilization rate of point card is a fraction that has the number of all accounting data as a denominator and the number of all accounting data including accounting data with point card as a numerator. The utilization rate of point card is a variable indicating the rate of users using parking on a regular basis because point card users most likely use parking regularly.

Table 1: Number of data for each cluster

cluster	1	2	3	4	5	6	7	8
number of data	19	126	152	116	153	150	150	184

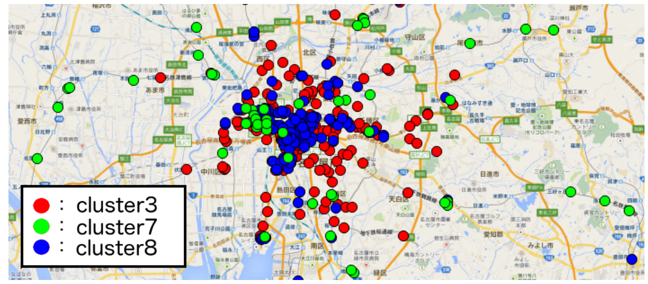

Figure 1: Cluster distribution

Discount rate of coupon The discount rate of coupon is a fraction that has the amount of total payment in all accounting data as a denominator and the amount of total discount with the coupon as a numerator. The discount rate of coupon is a variable indicating how active users discount parking charges.

Discount rate of maximum charge The discount rate of maximum charge is 1 subtracted by a fraction that has payment in case of not setting a maximum charge as a denominator and actual payment as a numerator. The discount rate of maximum charge is a variable indicating how effective users use a provided discount.

(2) Result of cluster analysis The number of parking spots for each cluster is presented in Table 1. In the spatial distribution feature of the clusters, the clusters are divided into a cluster that is located in a city center (urban types), a cluster that has a spatial wide distribution (suburban types), and a cluster that has a distribution feature between urban types and suburban types (middle types).

Cluster 2, cluster 6, and cluster 8 are urban types. Cluster 1, cluster 3, and cluster 5 are middle types. Cluster 4 and cluster 7 are suburban types. Distributions of cluster 3, cluster 7, and cluster 8 are illustrated in Figure 1. Points of red, green, and blue correspond to cluster 3, cluster 7, and cluster 8. Then, average and variance of distance of the nearest station are calculated for each cluster. As a result, difference in average distance of the nearest station is almost nothing. However, there are many parking lots around stations in cluster 1, cluster 4, and cluster 7 because these variances are low.

Characteristics of cluster 3, cluster 7, and cluster 8 are explained according to the above results and the figure illustrating 28 variables, which is composed of the utilization rate of parking and the parking time (Figures 2 and 3).

Cluster 3 This cluster's distribution is a middle type. The utilization rate of parking and the parking time have middle

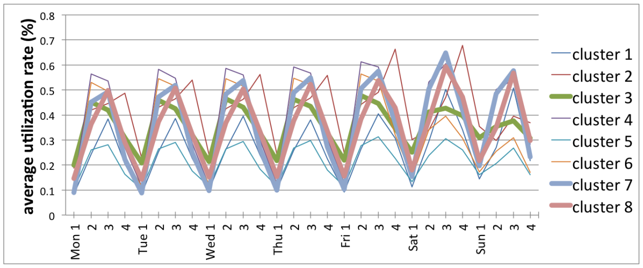

Figure 2: Average of utilization rate of parking for each cluster

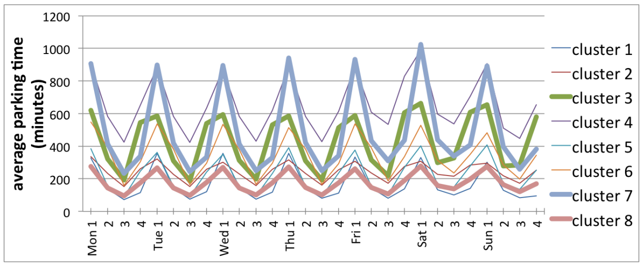

Figure 3: Average of parking time for each cluster

Characteristics; therefore, this cluster is a cluster to which general parking spots used for various applications belong.

Cluster 7 This cluster's distribution is a suburb type. Most of parking spots that belong to cluster 7 are located around train stations. The parking time between 1 o'clock and 7 o'clock is long. This cluster is used for a Park and Ride system. Motorists can park their cars near the station, and use trains to reach their destinations.

Cluster 8 This cluster's distribution is an urban type. The utilization rate of parking on holidays is high. The parking time is short. The discount rate of maximum charge is low. This cluster is one to which parking areas used for a short term belong.

2.3 Parking time model

Survival analysis (Klein and Moeschberger 2005) is a method that analyzes a relation between the time remaining until an event occurs and the event itself. Survival analysis is used for analyzing time until death by disease or until breakdown of a mechanical system in the fields of medical service, engineering, and so on. Time until an event occurs is called survival time. An analysis of relation between each covariate becomes possible by regarding parking time as survival time. A probability distribution is a weibull distribution (Weibull 1951) because parking time is drawn from weibull distribution (Kiyoshi 1968). From the above assumption, a probability of survival can be represented by the following formulation as in (1)

$$S(x|Z) = [exp(-\lambda x^{\alpha})]^{exp(\beta^T Z)} \quad (1)$$

where $S(x|Z)$ is the probability of survival given Z at survival time x, Z is a column vector of covariates, β^T is a row vector of coefficients, λ and α are parameters of weibull distribution. Influence of charge per hour, whether to set maximum charge, time zones in parking, holiday, and payment method as covariates on parking time is examined for each cluster by survival analysis. Survival analysis is executed for each cluster because the influence of each covariate on parking time by the characteristics of parking utilization differs. The covariates are explained below.

Charge per hour Value of charge per unit time is a variable of charge per hour. Not charge per unit time but charge per hour is used because unit time differs, such as 100 yen per 30 minutes or 100 yen per 90 minutes, in each parking spot.

Whether to set maximum charge If the maximum charge is set at the time a car enters a parking lot, the value of the variable is 1; otherwise, the value is 0.

Time zones in parking Time zones in parking are presented with three variables. Three variables are all 0, or one of the three is 1. Spans divided into four time zones in a day are assigned each variable in the same way as variables of the utilization rate for cluster analysis in 2.2.

Holiday If a day when a car enters a parking lot is a holiday, the value of the variable is 1; otherwise, the value is 0. A holiday indicates a Saturday, a Sunday or a festival day.

Payment method Three variables indicate whether to use a coupon ticket, a credit card, or a point card in charge payment. If a coupon ticket, a credit card, or a point card are used, the value of the variables is 1; otherwise, the value is 0.

As mentioned above, survival analysis is executed by a total of nine variables in charge per hour, whether to set maximum charge, time zones in parking, holiday, and payment method. If the value of each coefficient is positive, the coefficients have the probability of exiting a parking lot decreasing by survival analysis; otherwise, the coefficients have the probability increasing. The higher the absolute value of a coefficient is, the bigger the influence of a coefficient is. The values of coefficients are presented in Figures 4, 5, 6, and 7. The value of each covariate is considered in cluster 3, cluster 7, and cluster 8.

Setting price The coefficient of setting price means the coefficient of charge per hour, and the coefficient of whether to set maximum charge is considered. The coefficient of setting price is presented in Figure 4.

The absolute coefficient value of setting price is much lower than the other coefficients because the other coefficients are multiplied by the value of 0 or 1, while the coefficient of setting price is multiplied by the value of charge per hour such as 100 or 200. An axis of the coefficient of charge per hour is inverted because its value is negative. As for the coefficient of whether to set maximum charge, the influence is significant in cluster 3, and is not significant in cluster 7. As for the coefficient of charge per hour, the influence is significant in cluster 3, and is insignificant in cluster 7 and cluster 8. Cluster 3 is a cluster where change in parking time depending on charge revision is the most drastic because the absolute values of charge per hour and whether to a set maximum charge are the biggest. Cluster 7 is a cluster where charge in parking time depending on change of

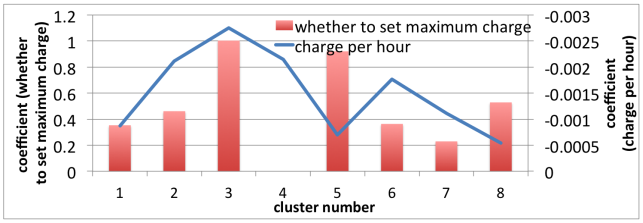

Figure 4: Values of charge per hour and whether to set maximum charge for each cluster

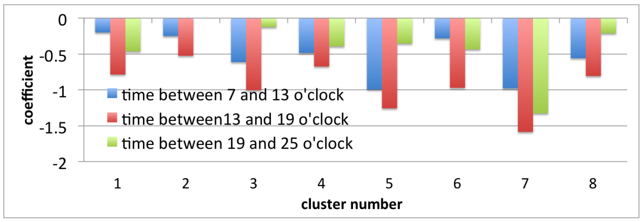

Figure 5: Value of time zones in parking for each cluster

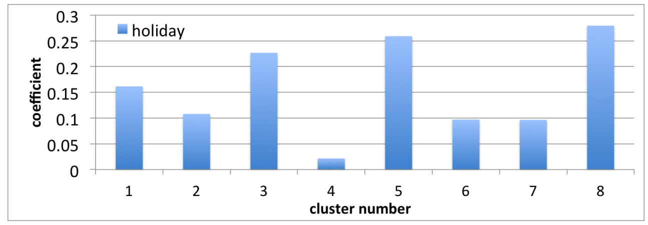

Figure 6: Value of holiday for each cluster

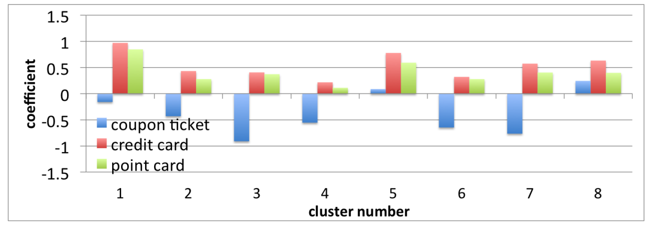

Figure 7: Value of payment method for each cluster

charge is small because the absolute values of charge per hour and whether to a set maximum charge are small.

Time zones in parking The coefficients indicating time zones in parking are presented in Figure 5. The coefficients are negative in all clusters. Therefore, users who park between 1 and 7 o'clock are users who park most probably for a long time. Conversely, users who park between 13 and 19 o'clock are users who park most probably for a short time because the absolute values of coefficients are the biggest. In cluster 7, parking time changes drastically depending on time zones in parking because the absolute values of the coefficients are big.

Holiday The coefficients about holidays are presented in Figure 6. Parking time becomes long on holidays because the coefficients are positive in all clusters. The influence of these coefficients is significant in cluster 3 and cluster 8.

Payment method The coefficients about payment method are presented in Figure 7. The values of the coefficients regarding credit card and point card are all positive, and the absolute values of these coefficients are mostly the same for each cluster. The probability of parking for a long time regarding users using credit cards or point cards becomes higher than the probability regarding users not using them. On the other hand, parking time regarding users using coupon tickets tends to become short except in cluster 8. Parking time tends to become long because the coefficients are positive in cluster 8. Cluster 8 can discount little parking charge from the maximum charge because users park for a short time in cluster 8. Therefore, there is willingness to park for a longer time than usual by using a coupon ticket and minimizing the parking charge. In parking lots where parking time is long, influence of a coupon ticket is significant, and parking time of users using coupon tickets becomes long. If parking time is long, the discount rate from the maximum charge tends to be high. Thus, incentive to minimize the parking charge by using a coupon ticket becomes low. Cluster 3 is a cluster where parking time becomes short by using a coupon ticket, and it is the most significant of all the clusters. The parking time of cluster 3 is average length. Influence of charge per hour and whether to set maximum charge in cluster 3 is the most significant of all the clusters. Thus, incentive to minimize parking charge is the highest, and parking time of users using a coupon ticket becomes short.

The influence of each covariate for cluster 3, cluster 7, and cluster 8 on the parking time is summarized from the above examination.

Cluster 3 The influence of setting price, holiday, and using coupon ticket is significant. Prudence is necessary for setting price because cluster 3 is a cluster where willingness to minimize parking charge is high.

Cluster 7 The influence of setting price is insignificant. The influence of time zones in parking is significant. Changing the utilization status is possible by setting different prices depending on time zones, for example, setting price lower except between 1 and 7 o'clock.

Cluster 8 The influence of charge per hour is not so significant. The influence of holidays is significant. Increasing revenue is possible by increasing charge per hour on holidays because parking time becomes longer on holidays.

2.4 Change points of utilization characteristics

The characteristics of parking utilization are changed by season or surrounding environment. Therefore, detecting change points of utilization characteristics is important for setting parking fee. In this study, we try detecting change points by assigning utilization rate and parking time per day to each typical one with cluster analysis, and considering changes in these as changes in clusters. Detecting the change points of the utilization characteristics can become easy by classifying the utilization characteristics and parking time per day, which is difficult to understand by just checking raw data, into clusters because the utilization characteristics

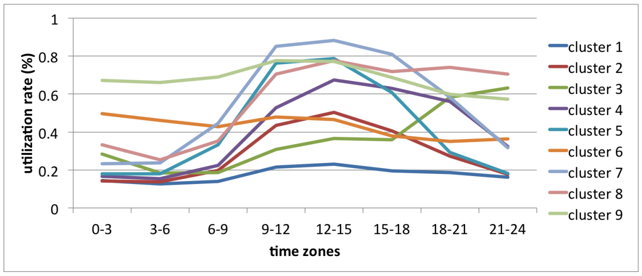

Figure 8: Cluster centers of utilization rate per day

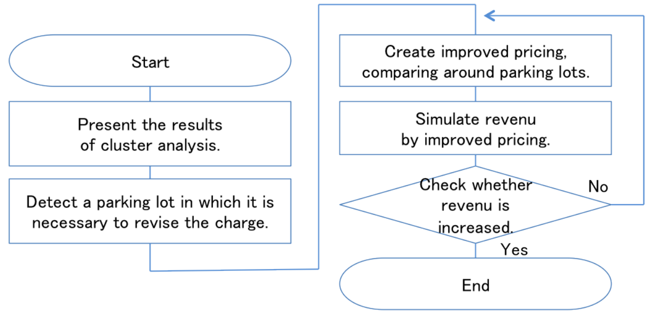

Figure 9: Flowchart of pricing procedure

are presented as one value. The utilization rate and the parking time are calculated in each time zone dividing a day into 8 zones of 3 hours each from 0 o'clock, and cluster analysis is executed regarding changes in each utilization rate and parking time as 8 variables by data for 365 days on the 1,050 parking lots.

The cluster number of the utilization rate is 9. The cluster number of the parking time is 7. The cluster centers of the utilization rate are presented in Figure 8. Cluster numbers are assigned in an ascending order of summing each value in cluster centers. In the next section, we consider whether influence of charge revision causes a change leading to increased revenue by changes in clusters, and the examination is used for setting price. (e.g., Figure 12 in Section 3.2).

3 Price procedure based on data

3.1 Summary of setting price and simulation based on data

In the section, we propose a pricing procedure. First, we present the results of cluster analysis in 2.2 on Google Earth, and detected a parking lot in which it is necessary to revise the charge. These results are presented on Google Earth because the surrounding environment can be easily understood. The parking clusters existing in great number are adapted for users' demands because users' demand for nearby parking is similar. A parking belonging to a parking cluster differing from parking clusters existing in great number are not adapted for users' demands. Thus, if a parking lot belonging to a parking cluster differing from a parking cluster of great number exists in some regions, the parking lot is regarded as needing to revise the charge. Second, better pricing is considered by examining the difference between the parking lot needing to revise its charge and surrounding parking lots. As for a parking cluster to which parking lots needing to revise their charge belong, a parking lot increasing revenue by charge revision is detected by checking change of clusters regarding the utilization rate and parking time made in 2.4, and the results are used for setting price in the parking lot needing to revise its charge. Finally, whether improved pricing is appropriate is considered by simulating how long the parking time charges by setting price with survival analysis generated in 2.3, and considering influence on revenue. A flowchart of the above pricing procedure is presented in Figure 9.

Figure 10: Parking lots around Hisaya-odori station

3.2 Executing setting price based on data

The parking lots around Hisaya-odori station are presented in Figure 10. The pink pins mean the lots belonging to parking cluster 2. The orange pins mean the lots belonging to parking cluster 6. The white pin means the lot belonging to parking cluster 8. The parking lot belonging to parking cluster 8 is indicated by a red arrow in Figure 10. Therefore, parking lot 1 needs to revise its charge. The improved pricing is considered by comparing the pricings or the changes of utilization rate clusters in the nearest parking lot 2 along the same road.

The changes of the utilization rate clusters in parking lot 1 and parking lot 2 are presented in Figure 11. There are many utilization rate cluster 2 and utilization rate cluster 4 in Figure 11 regarding parking lot 1. Conversely, there are many utilization rate cluster 4 and utilization rate cluster 7 in Figure 11 regarding parking lot 2, thus the utilization rate is high. The pricings in parking lot 1 and parking lot 2 are presented in Table 2. Comparing pricing in parking lots 1 and 2, the charges per unit time are mostly the same; however, the settings of maximum charge are very different. Thus, setting no maximum charge in the parking lot 1 between 8 and 24 o'clock causes a large decrease in the utilization rate. Figure 12 shows the changes of the utilization rate clusters in the parking lot belonging to the same parking cluster 8 as lot 1. In Figure 12, there are many days assigned the utilization rate cluster 1 and utilization cluster 2 because the utilization rate is low before the 310th day; however, there are some days assigned the utilization rate cluster 7 and utilization rate cluster 8 because the utilization rate is increas-

Table 2: Pricing in parking lot 1 and parking lot 2

	charge type	day type	time zones	unit time	unit charge
1	unit time	all day	0 - 8	60 min	100 yen
	unit time	all day	8 - 24	30 min	200 yen
	maximum charge	all day	0 - 8	-	500 yen
2	unit time	all day	0 - 24	30 min	200 yen
	maximum charge	all day	22 - 8	-	500 yen
	maximum charge	weekday	8 - 22	-	1500 yen
	maximum charge	sat, sun, holidays	8 - 22	-	1000 yen

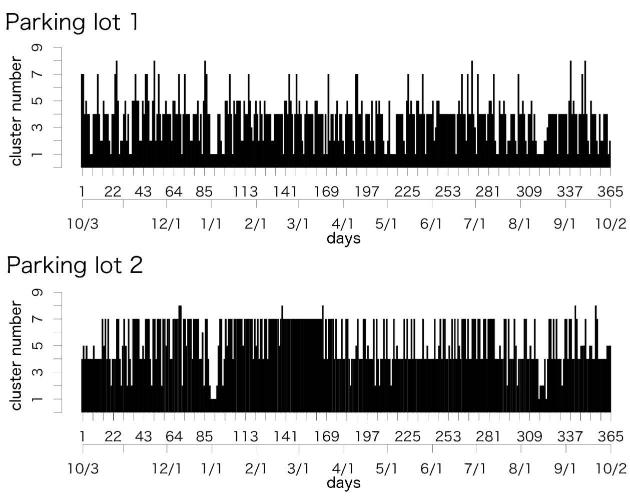

Figure 11: Changes of utilization rate clusters in parking lot 1 and parking lot 2

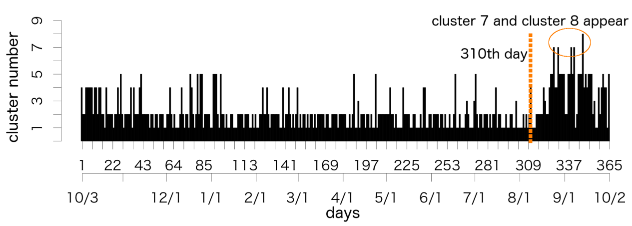

Figure 12: Changes of utilization rate clusters in lot belonging to parking cluster 8

ing by setting a maximum charge near the 310th day. Setting maximum charge is effective for increasing parking time because influence of maximum charge is more significant than charge per hour in parking cluster 8. We propose pricing that sets 1,500 yen in maximum charge between 8 and 24 o'clock from the above consideration as better pricing in parking lot 1. After this, this proposed pricing is called improved pricing 1. We consider appropriateness of this pricing in the next section.

3.3 Evaluation of price with simulation

(1) Simulation method The appropriateness of the pricing is evaluated by estimating parking time with survival analysis, and calculating revenue.

1. A accounting data in a target parking lot is retrieved.
2. The covariates except charge per hour and whether to set maximum charge from the retrieved data are decided .
3. The values of charge per hour and whether to set maximum charge is decided based on considering pricing, and a survival function $S(x)$ is constructed.
4. A value between 0 and 1 is generated randomly, and the value is defined as p.
5. x is incremented by 10 from 10 until $S(x) < p$ is corrected, then parking charge is calculated and stored regarding x satisfying the conditions as parking time.
6. Procedures 1 to 5 are repeated until all target accounting data are retrieved.

The value of x meaning the parking time is incremented by 10 from 10 because a minimum value of unit time in charge per unit time is 10 minutes. However, maximum parking time is 4,320 minutes, equal to 3 days, in order not to generate too long a parking time.

(2) Simulation result Figure 13 shows a distribution of the actual parking time and a distribution of the parking time estimated by not changing pricing. Figure 14 shows a distribution of the parking time estimated by not changing pricing and a distribution of the parking time estimated by the improved pricing 1. The parking time is estimated every 10 minutes in the simulation; however, the users are counted every 30 minutes from 0 minutes in Figures 13 and 14 because the minimum value of unit time in charge per unit time is 30 minutes. The horizontal axis indicates the parking time, and is shown until 900 minutes when the number of users approaches 0. Table 3, Table 4, and Table 5 show average parking time, the number of overflowing users, total revenue, and charge per person in actual parking time, the parking time estimated by not changing pricing, and the parking time estimated by changing pricing. Charge per person indicates a value of dividing the total revenue by the number of users. The number of overflowing users indicates the number of users who can't park because the parking lot is full. The longer the parking time becomes, the more the number of overflowing users becomes. The number of overflowing users needs to be considered because these become the factor decreasing the revenue in parking lots where maximum charge is set. Thus, two values of whether to consider the number of overflowing users are presented regarding the

Table 3: Comparison of the average parking time and the number of overflowing users between actual parking time and estimated parking time

	the average parking time (minutes)	the number of overflowing users
the actual data	93.35	-
the data estimated by not changing pricing	85.33	2,437
the data estimated by changing pricing	150.20	6,957

Table 4: Comparison of the total revenue between actual parking time and estimated parking time

	the total revenue (yen)	
	A	B
the actual data	17,246,400	-
the data estimated by not changing pricing	17,224,500	15,886,500
the data estimated by changing pricing	23,938,400	18,705,300

Table 5: Comparison of the charge per person between actual parking time and estimated parking time

	the charge per person (yen)	
	A	B
the actual data	570.96	-
the data estimated by not changing pricing	570.23	572.09
the data estimated by changing pricing	804.56	792.50

Table 6: Simulation result in case of setting excessively inexpensive maximum charge

the total revenue (yen)		the charge per person (yen)	
A	B	A	B
20,031,600	15,661,900	663.17	673.66

total revenue and the charge per person in Table 4 and Table 5. The value of not considering the number of overflowing users is indicated by A; in contrast, B indicates the value of considering the number of overflowing users. In the case of not considering the number of overflowing users, a total revenue is calculated by all users even if the parking lot is full. In the case of considering the number of overflowing users, a total revenue is calculated by the users who can park.

Accuracy of the simulation is considered comparing the actual parking time and the parking time estimated by not changing pricing. A time gap is approximately 8 minutes comparing the average parking time in the actual data and the data estimated by not changing pricing in Table 3, and the total revenue and the charge per person can be mostly approximated in Table 4 and Table 5. Comparing the distributions of the parking time in Figure 13, the distributions can be mostly approximated because there are the greatest number of users between 0 and 30 minutes, and the distributions decrease the number of users over time together in both the actual distribution and the distribution estimated by not changing pricing. However, the difference between the number of users between 0 and 30 minutes and the number of users between 30 and 60 minutes is large. Thus, the distributions need to be improved, and this is an issue that remains for the future.

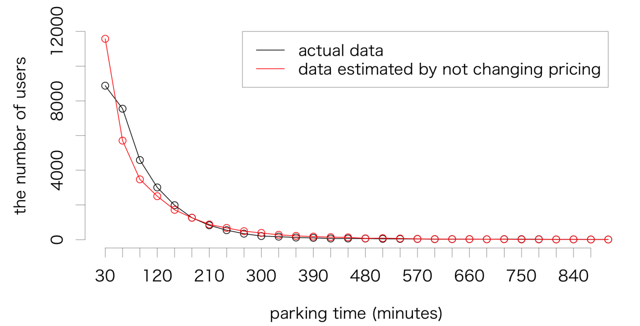

Figure 13: Distribution of actual parking time and distribution of parking time estimated by not changing pricing

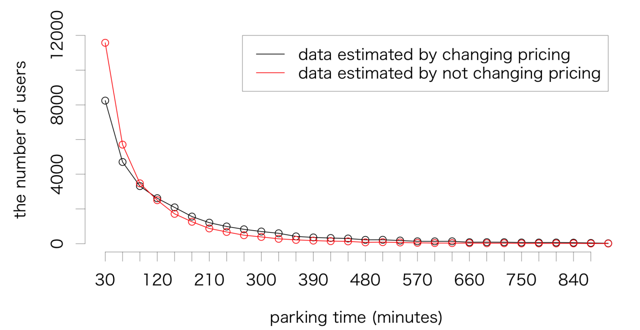

Figure 14: Distribution of parking time estimated by not changing pricing and by changing pricing

The appropriateness of the improved pricing 1 is considered. Comparing the distributions of parking time in Figure 14, the parking time becomes long by setting the maximum charge as the improved pricing. An increase in the utilization rate is expected by lengthening the parking time. However, if pricing becomes excessively inexpensive in order to increase the utilization rate, parking spaces tend to fill up by lengthening the parking time, and then the revenue decreases because the number of overflowing users increases. Therefore, pricing that doesn't make the turnover rate too low is needed. Comparing the data estimated by not changing pricing and by changing pricing, the average parking time became longer for approximately 65 minutes in the data estimated by not changing pricing. In the data estimated by changing pricing, the charge per person not considering the number of overflowing users (the charge per person A) and the charge per person considering the number of overflowing users (the charge per person B) increase approximately 230 yen and 220 yen in comparison with the data estimated by not changing pricing because of increasing the parking time. The total revenue not considering the number of overflowing users (the total revenue A) increases approximately 6.7 million yen, and the total revenue considering the number of overflowing users (the total revenue B) also increases approximately 2.8 million yen.

Table 6 shows the simulation result by pricing that sets 1,000 yen in maximum charge of parking lot 1 between 8 and 24 o'clock in Table 2. The parking time is not changed by changing the value of maximum charge; then the average parking time and the number of overflowing users don't increase because just whether to set maximum charge is a covariate in the generated parking time model in this study. Thus, the average parking time and the number of overflowing users are the same values as the data estimated by changing pricing in Table 3. In Table 6, the total revenue not considering the number of overflowing users (the total revenue A), the charge per person not considering the number of overflowing users (the charge per person A) and the charge per person considering the number of overflowing users (the charge per person B) increase comparing the data estimated by not changing pricing in Table 4 and Table 5; however, the total revenue considering the number of overflowing users (the total revenue B) decreases by approximately 0.2 million yen. In addition, the total revenue A, the total revenue B, the charge per person A, and the charge per person B decrease comparing the data estimated by changing pricing. Therefore, 1,000 yen is excessively inexpensive regarding maximum charge. The improved pricing setting 1,500 yen as maximum charge is an appropriate pricing for increasing revenue.

4 Related work

Asano et al. (Yoshihiro 2007) suggests that low revenue by low turnover rate in public parking spaces is a problem. The purpose of exploring use of flexible demand based parking charges is to address problem of low revenue as noted. This paper proposes a method for maximizing total revenue. Under the method, the charge becomes high when the number of users that enter the parking lot is many and utilization rate is high, while the charge becomes low when the number of users that enter the parking lot is few and utilization rate is low. As an evaluation of the dynamic pricing method, total revenue is calculated by simulation based on actual data for the number of users.

Hashimoto et al. (So, Ryo, and Takayuki 2013) suggests an effective parking reservation system, assigning parking spaces and deciding parking charge by auction. Users who have high willingness to pay are assigned parking spaces by auction, which determines the parking charge in accordance with uesrs' willingness to pay. From this, an increase in users' convenience and parking revenue is expected. As an evaluation of the proposed reservation system, a simulation considering price elasticity is executed. For this simulation, clusters are composed by cluster analysis from actual data for parking utilization, and parking time models are created for each composed cluster by survival analysis.

Ishigaki et al. (Tsukasa, Takeshi, and Yoichi 2011) suggests that low productivity is a problem because the quality of service industry depends on service provider's experience and intuition. This paper aims to improve profitability by coupon strategies or low price wars, focusing on retail business in service industries. The purpose of this paper is supplying a system or method from large scale data about past daily buying behavior. Under the system, a non-expert can systematically create categories of customers and products that have potential for an effective promotion. In concrete terms, customer behavior can be understood by constructing a Bayesian network based on categories after extracting potential customers and product categories from questionnaire data and ID-POS data.

We now explain the position of this study. The purpose of this study is to enable a non-expert to set parking charges adapted for the utilization characteristics by systematically extracting parking categories from large-scale data about parking accounting such as (Tsukasa, Takeshi, and Yoichi 2011). As a concrete method, we classify characteristics of parking utilization into some clusters by cluster analysis, then we create parking time models for each cluster by survival analysis such as (So, Ryo, and Takayuki 2013). Unlike (So, Ryo, and Takayuki 2013), we execute analysis in more detail by actual pricing change data in addition to parking accounting data. We propose a more appropriate pricing from the result of cluster analysis and survival analysis, then we consider appropriateness of the pricing by simulating total revenue such as (Yoshihiro 2007).

5 Conclusion

In this study, a pricing procedure based on data analysis was proposed using parking accounting data on 1,050 parking places around Nagoya City over recent one year. First, the parking lots were classified according to utilization characteristics by cluster analysis in order to discuss the characteristics of the parking utilization. Influence on parking time, which occurred by a difference of covariates such as the pricing and the time zones when users parked, was analyzed for each cluster by survival analysis after the characteristics of each cluster were considered. The characteristics of the parking utilization for a year were analyzed by the utilization characteristics by cluster analysis, and the analysis of the influence on the parking time by survival analysis. However, the utilization characteristics changed over time. Therefore, the dynamic utilization characteristics and the parking time per day were classified into clusters by another cluster analysis. Detecting the change points of the utilization characteristics became possible by considering the change of the clusters. Considering appropriateness of pricing became possible by simulation. The proposed pricing procedure enables a non-expert to set changes adapted for the characteristics of the parking utilization.

Acknowledgement

The parking accounting data was provided by Meitetsu Kyosho Corp. We would like to express our gratitude.

References

Kiyoshi, K. 1968. The study on actual state of parking in service area of expressway (part 2) : Distribution of parking durations. *SEISAN KENKYU* 20(7):362–364.

Klein, J. P., and Moeschberger, M. L. 2005. Survival analysis: Techniques for censored and truncated data (statistics for biology and health).

So, H.; Ryo, K.; and Takayuki, I. 2013. A simulation evaluation of auction-based parking reservation system based on parking utilization data. *The Special Interest Group Technical Reports of IPSJ* 2013(23):1–7.

Tsukasa, I.; Takeshi, T.; and Yoichi, M. 2011. Customer behavior prediction system by large scale data fusion in a retail service. *The Japanese Society for Artificial Intelligence Journal* 26(6):670–681.

Weibull, W. 1951. A statistical distribution function of wide applicability. *Journal of Applied Mechanics* 18:293–297.

Yoshihiro, A. 2007. A study of pricing in public parking space. Master's thesis, Nanzan University graduate schools Mathematical Sciences and Information Engineering.