TRR
JOURNAL OF THE TRANSPORTATION RESEARCH BOARD

*Article*

Transportation Research Record
1–12

DOI: 10.1177/0361198118772947
journals.sagepub.com/home/trr
SAGE

# Predicting Owners' Willingness to Share Private Residential Parking Spots

**Chu Zhang[1], Jun Chen[1], Zhibin Li[1], and Yuanyuan Wu[2]**


## Abstract

Sharing of private residential parking spots is a new pattern of parking management in China. This pattern corresponds to the booming sharing economy and is growing very fast. It can significantly improve the utilization of parking resources and relieve parking supply pressure. Based on the real data of 1-year behavioral records of owners obtained from Ding Ding Parking (DParking), an application on smart phones, as well as various field survey data, the study analyzed the influential factors and predicted owners' sharing willingness. Two Classification and Regression Trees (CART) were developed to answer questions pertaining to whether owners would share their parking spot and how long owners would share during peak periods of parking demand, respectively. The results showed good accuracy in both models and revealed that owners' self-use behavior, along with owners' private spots' physical characteristics and rental effects of the previous month, all have significant influence on owners' willingness to share. The influence of factors and their importance differ for the two models; thus, a detailed comparison is performed. The findings in this paper would be beneficial to the government's parking supply policies, as well as to third parties, so as to enhance the effective distribution of parking resources.


Private residential parking spot sharing has recently become a hot topic in China. Generally, parking demands are mainly centered around places of work in the daytime and then transfer to residential zones by early evening or night (*1*). Making private residential parking spots accessible would promote the utilization of parking resources and be more compatible with both owners' and parkers' interests (*2*). Because of the legal issues admitted by the government and the technologies developed for parking management, owners of private parking spots now have access to share spare parking resources with the public, in exchange for economic payback. Because the ownership of parking spots in residential areas tends to be scattered, transactions facilitated through an app-based peer-to-peer marketplace were adopted instead of the more traditional parking features, such as parking meters, gate machines, and so on. Some firms take on the role as a third party, so as to collect and distribute parking resources through the use of various smart phone applications, such as ROVER in the United States and DParking in China. The detailed operation steps of both owners and borrowers adopting DParking are listed in Figure 1. Thus, the new pattern of parking supply generates numerous fluctuations in shared parking studies.

Firstly, shared parking behavior includes the participation of both the owners and parkers. Current studies have already taken many factors into consideration, such as parking fees, walking distances, trip purpose, parkers' psychology, and so on (*3, 4*). However, all the studies above only considered the participation willingness of parkers, and none of the studies has explored aspects concerning the sharing willingness of owners.

Moreover, the traditional research method for willingness primarily focuses on stated preference surveys (*5, 6*); the shortcomings are obvious. The accuracy of the answers mostly depends on the quality of the questionnaires. Inevitably, vague descriptions of the preset context may tend to mislead respondents, which may result in wrong or skewed answers. Thanks to the new technologies, the revealed preference survey can now be conducted through the behavioral record documents saved in the database; thus, a real-data based study would prove to be more persuasive.

Another difference is the level of study objects. Traditional studies are based on the parking lots (*3, 7,*

[1]School of Transportation, Southeast University, Nanjing, China
[2]School of Civil and Environmental Engineering, Nanyang Technological University, Singapore

**Corresponding Author:**
Address correspondence to Jun Chen: chenjun@seu.edu.cn

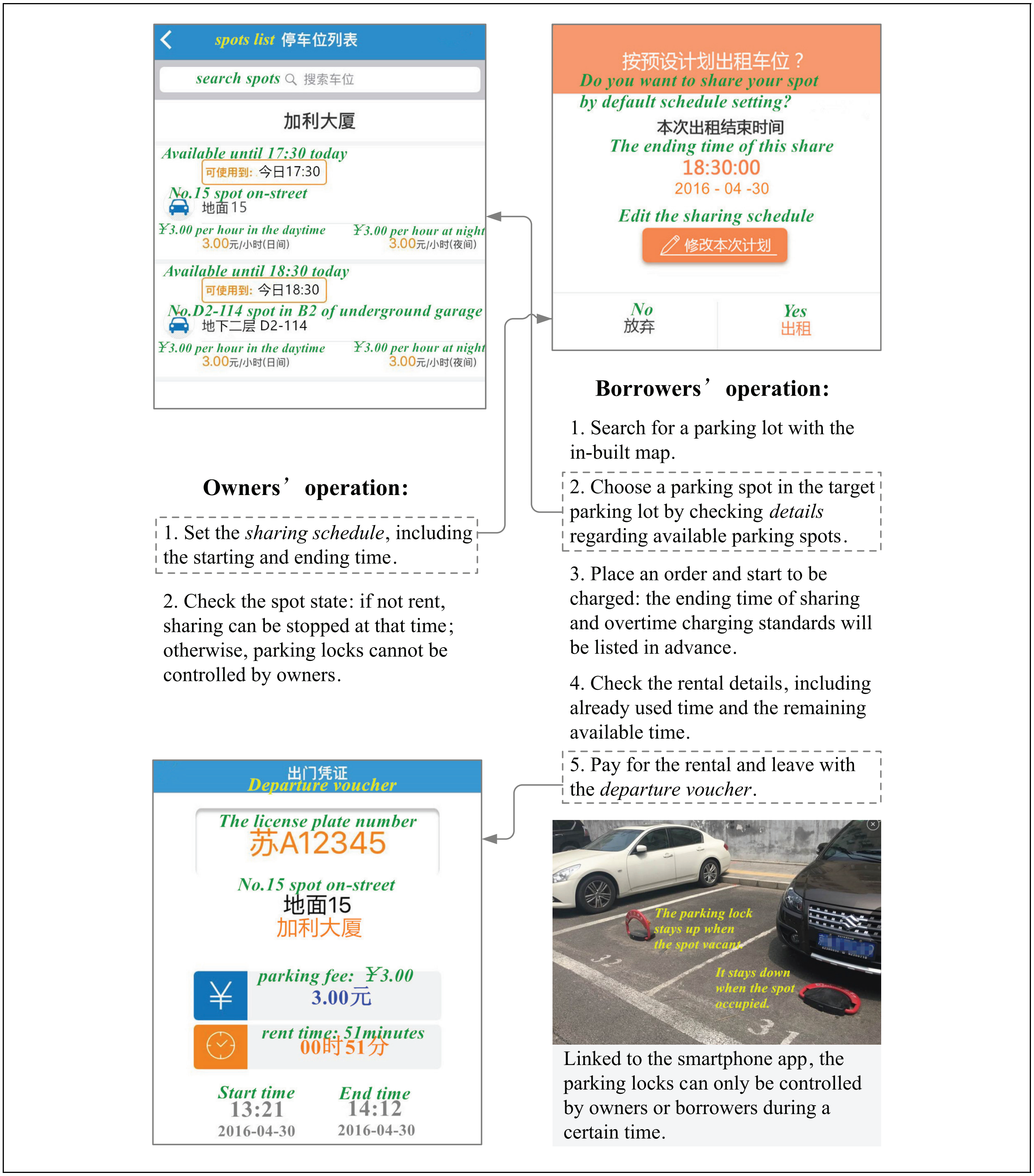


**Figure 1.** Operation of DParking app and parking locks.

8), whereas, in this study, every spot is assumed to be associated with its owner and to be different from others. Spot-level research is seldom up-to-date and mostly focuses on the distribution methodology (*9, 10*), without any adequate real-data support. Real study data encompass detailed information about both the parking spot's physical characteristics and its owner's behavioral records, perhaps for the first time.

As a result, this paper conducts a spot-level study of owners' sharing willingness driven by real-data for the first time. The primary objective of the research is to accurately investigate factors affecting the sharing willingness of private parking spots' owners. More specifically, the study presented in this paper focuses on answering the following two questions: (1) would owners decide to share or not? (2) how long would owners share their parking spots during the peak periods of parking demand?

## Data

### Study Area

The research was conducted in a residential building on the north side of the 2nd Ring Road of Beijing, adjacent to two office buildings and one government building. The target building consists of a two-level underground parking garage as well as a few on-street parking spots. As depicted in Figure 2, the on-street parking zone stretches for less than 200 m, and lies along the alley across from major buildings. The underground parking garage also has an entrance and an exit for automobiles separately along the parking zone. The longest distance between any spot and the furthest surrounding building is 247 m (810 ft), which can be accepted by most parkers (*11*).

The procedure for shared parking is as follows. Owners post the open-for-share information through the app, and ensure that the parking spots are available during the preset period. If the spots are not booked or rented, the owners can withdraw or modify their sharing plans at any time. Parking prices are fixed at 5 yuan (0.8 USD) per hour throughout the daytime period from 9:00 am to 19:00 pm and 1 yuan (0.2 USD) per hour for the remaining time. The automatic charge unit is per 15 min., and the charging scheme is the same on workdays and weekends.

### Data Processing

Study data were obtained from DParking, an application on smart phones, and were truncated from all continuous operation records for a period of an entire year, starting at 2015/11/01 00:00:00 and ending at 2016/10/31 23:59:59. Because seasonal heterogeneity influences the travel behavior of both the owners and renters, the samples were aggregated per spot per month. Therefore, a total of 604 samples were assembled among 72 spots representing various types of owners. The statistical results revealed that the total rental time per day on workdays was 2.35 times that of weekends. Moreover, it revealed the total rental time per hour in the daytime (from 7:00 am to 19:00 pm) was 2.56 times that of night rentals. Thus, willingness to share in the daytime on workdays is crucial to meet the growing demands and promote utilization. The period from 7:00 am to 19:00 pm on workdays was defined as the peak time of parking demand in this study.

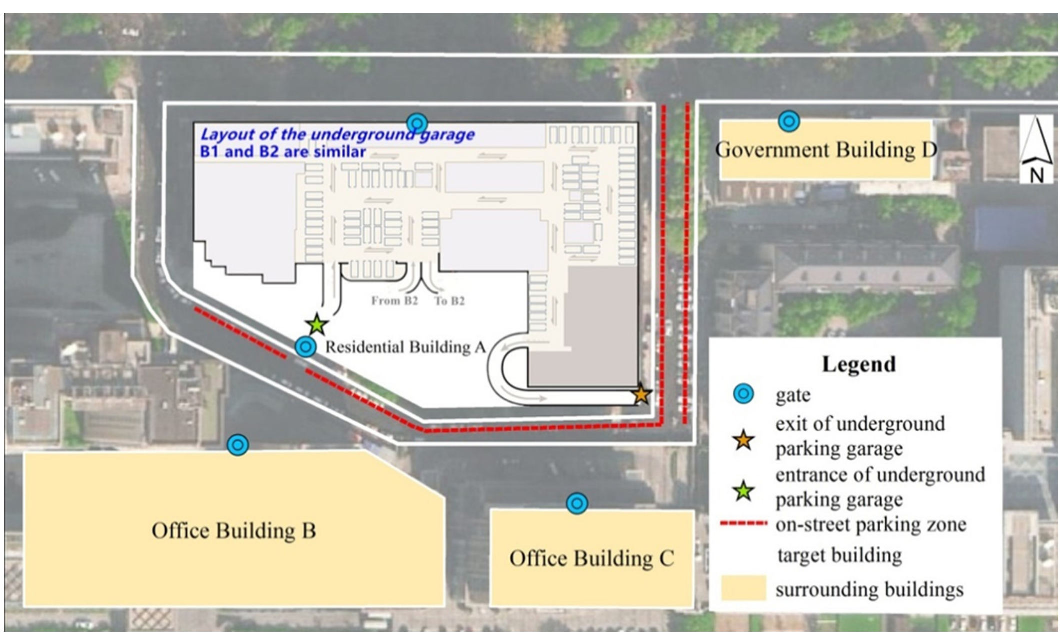


**Figure 2.** Layout of the study area.

**Table 1.** Descriptions of Variables

| Variables | Descriptions | Avg | SD | Avg | SD |
|---|---|---|---|---|---|
| | Owners' sharing decisions | | | | |
| osYN | Open-to-share decision | 0: Unshare (227 samples) | | 1: Share (377 samples) | |
| osDtime | Average sharing time in the daytime on workdays (h) | na | | 5.74 | 3.54 |
| | Spots' physical characteristics | | | | |
| Type | 0: underground; 1: on-street | 0.50 | 0.50 | 0.50 | 0.50 |
| Floor | 1: on-street; 2: B1 floor; 3: B2 floor | 1.68 | 0.76 | 1.76 | 0.84 |
| distanceL | Distances from spots to the nearest surrounding building (m) | 80.17 | 41.70 | 71.01 | 42.98 |
| distanceA | Average distances from spots to surrounding buildings (m) | 133.37 | 47.17 | 130.23 | 44.34 |
| | Owners' self-use behavior | | | | |
| sfTtime | Average self-use time in a day (h) | 16.14 | 6.11 | 10.17 | 7.79 |
| sfDtime | Average self-use time in the daytime on workdays (h) | 8.06 | 3.13 | 3.94 | 3.50 |
| sfDfres | Average self-use parking frequency in the daytime on workdays | 2.31 | 0.23 | 1.95 | 0.41 |
| | Rental effects of the previous month | | | | |
| rtTtime | Average rental time in a day (h) | na | | 2.58 | 4.04 |
| rtDtime | Average rental time in the daytime on workdays (h) | | | 2.29 | 2.72 |
| rtosDR | Ratio of rental time in sharing time in the daytime on workdays | | | 0.52 | 0.29 |

*Note*: Avg = average; SD = standard deviation.

The entire study data include 377 shared samples and 227 unshared samples. Variables of the two types of samples are described separately. All 12 variables were collected and organized into four specific groups (see Table 1), including the owners' sharing decisions, the parking spots' physical characteristics, owners' self-use behavior, and the rental effects of the previous month. The owners' sharing behavior group consists of two dependent variables: osYN and osDtime; osYN is a binary variable indicating whether the sample is available for sharing or not. Additionally, osDtime, as a continuous variable only existing in the shared samples, is the average time of sharing behavior in the daytime during workdays. In the spots' physical characteristics group, Type and Floor constitute the fundamental factors, distanceL describes the minimum distances between the spot and the gate of the nearest surrounding building, and distanceA represents the average value of the distances from the spots to all the gates of the three surrounding buildings. In the owners' self-use behavior group, sfTtime is the average value of the owners' self-parking time in a day, whereas sfDtime and sfDfres depict the average time and frequency owners parked in the daytime on workdays. As for rental effects of the previous month's group, all the variables were collected from records for the previous month, the one before the statistical month, as owners' prior knowledge, which may have influenced the current month's sharing decisions. Moreover, rtTtime and rtDtime are the average rental time in a day and in the daytime on workdays, respectively. The rtosDR variable depicts the ratio of rental time in sharing time in the daytime on workdays. Like osDtime, these three variables also only exist in shared samples.

## Method

Most sharing willingness studies have adopted parametric statistical approaches, such as various logit models (*2–4*). However, some researchers also built nonparametric models which can provide higher prediction accuracy like using CART (*12, 13*). CART is a typical machine learning method and has several advantages perfectly matching this study's targets. First, there are two kinds of CARTs, classification tree (CT) and regression tree (RT), distinguished by the types of target value. As in this case, one dependent variable, osYN, is dichotomous and the other, osDtime, is continuous, both CT and RT are required. Considering that the influential factors in both models are similar, adopting both CART models would benefit the comparison. Secondly, CART reflects the complex nonlinear correlation between the influencing factors and target behavior variables and, thus, produces "if–then" rules (determining the willingness according to the conditions), so as to predict owners' sharing choices. Besides, CART can graphically depict the correlation, which aids in understanding how owners make their decisions.

Although the leaf value of CT is represented by the majority and that of RT is represented by the average, the development of a CART model is the same. It generally consists of three steps: (1) growing the tree, (2) pruning the tree, and (3) selecting an optimal tree from the pruned trees. The detailed procedure relevant to this paper is as follows:

Step 1: To grow the tree is to recursively partition the target variable, so that the data in the descendent nodes are always more pure than the data in the parent node. The splitting methods are differentiated between CT and RT. For CT, node impurity based on entropy reduction was adopted, as it would be a stricter splitting criterion for a complex and chaotic set than the traditional Gini index (*14*). For RT, on the other hand, node impurity based on the least squares method was adopted.

For a given node $k$ in CT, the node impurity $i(k)$ is calculated

$$i(k) = -\sum_{j=1}^{2} p_j \log_2 p_j \tag{1}$$

where $p_j$ is the proportion of class $j$ in node $k$.

For a given node $k$ in RT, the node impurity $i(k)$ is calculated as

$$i(k) = \frac{1}{N}\sum_{i=1}^{N} (y_i - r_i)^2 \tag{2}$$

where $N$ is the total number of samples in node $k$; and $r_i$ represents the average value of dependent variable $y_i$.

Moreover, for all the possible splits in the values of the explanatory variable $x$, the impurity reduction on the parent node $k$, caused by a split $s$ is calculated as

$$\Delta i(x, s, k) = i(k) - p_L i(k_L) - p_R i(k_R) \tag{3}$$

where $k_L$ and $k_R$ are the left and right child nodes of the parent node $k$, respectively; $p_L$ and $p_R$ are fractions of the records belonging to the left and right child nodes, respectively. If $\Delta i(x, s, k) > 0$, this implies that the records in the descendent nodes are purer than those in the parent node. If not, then the split is not viable. The split scheme contributing to a global maximum impurity reduction constitutes the most appropriate split.

In this paper, splitting ceases when one of the following stopping rules is satisfied: (1) none of the nodes at the bottom can be split further when $\Delta i(x, s, k)$ is less than the allowed thresholds, which is 0 for CT and 0.01 for RT in this case; (2) records in the node are less than the preset minimum number, which is 4 in this case; and (3) the current tree depth has reached the preset maximum tree depth, which is 8 in this case.

Step 2: Tree pruning is used in case the tree is overfitted. The overfitted tree could result in a substantial true error rate on the new (testing) data (*14, 15*). Therefore, there is a critical need to remove some branches that do not contribute to predictive accuracy. Generally, the cost complexity measurement approach, based on the postpruning concept, is employed to prune the grown tree. The cost complexity of a tree $T$ is then defined as the weighted sum of its complexity (i.e., number of leaf nodes) and its error, expressed as

$$R_\alpha(T) = R(T) + \alpha |\widetilde{T}| \tag{4}$$

where $\alpha$ is the weight; $|\widetilde{T}|$ is the number of leaf nodes in tree $T$; and $R(T)$ is the corresponding misclassification error rate, which is node impurity in this particular case.

For a given node $t$ in the original tree $T_0$, tree $t$ also has its child tree $T_t$. The reduction index of total loss in pruned trees is expressed as

$$g(t) = \frac{R(t) - R(T_t)}{|\widetilde{T}_t - 1|} \tag{5}$$

The tree pruning process is recursively performed by determining the minimum $g(t)$ and pruning the corresponding child tree until only the root node remains. The optimal tree is in the set of pruned trees and is gained by each pruning.

Step 3: To select an optimal tree from the pruned trees is to calculate the cost complexity, based on the average values of global accuracy, or $R$-squared through the test data and select the one with a minimum value. In this paper, a 10-fold cross validation method was adopted for testing.

Moreover, by determining the final tree, the importance of the variables that intervene in the model can be derived from the CART model. This is the average reduction of entropy or $R$-squared in all the nodes. The value of the standardized importance of the independent variables reflects the impact of such predictor variables on the model (*16*).

## Results

### *CART_1: Whether to Share or Not?*

Because the unshared samples account for a small proportion of 37.6%, in order to eliminate biased results caused by unbalanced data, data preprocessing should be performed before building CART_1. Generally, the oversampling method for expanding minority data may increase the probability of the overfitting problem compared with the undersampling method for reducing the majority data (*17*). Therefore, the random undersampling method to reduce the shared samples is applied in this study. The input data include 227 unshared samples and 227 shared samples, as the share possibilities are assumed to be equal. As for the independent variables, factors in spots' physical characteristics and owners' self-use behavior groups were included in CART_1. The dependent variable here is the categorical variable, osYN.

Based on the preset parameters (illustrated in Methods) and the 10-fold cross validation testing method, the pruned CART_1 that leads to the lowest cost complexity for the testing data was determined. Figure 3 depicts the relationship between the number of leaves and cost complexity in CART_1; cost complexity reached the bottom at 11 leaves with a value of 0.167.

Figure 4 illustrates CART_1 with 11 leaf nodes; the shaded rectangle represents one of the leaf nodes. CART_1 can be intuitively interpreted as follows: the initial split at Node 1 (the root node) is based on sfDtime. This means the most important factor affecting owners' sharing decisions is the owners' self-use of time during the daytime on workdays. CART_1 directs values of sfDtime lower than 1.90 hours to the left, forming LeafNode 1, and values larger than 1.90 hours to the right, forming Node 2. As indicated by LeafNode 1, CART_1 predicts that owners parking their own vehicles for less than 1.90 hours during the daytime on workdays will probably be willing to share. Further, CART_1 proceeds to split Node 2, in accordance with sfTtime. If sfTtime is more than 21.96 hours a day, the not-to-share decision is most likely to be made (97.2%), as is depicted in LeafNode 11. This is because the self-use parking has already occupied the whole day and the available time is too limited. When sfDtime exceeds 1.90 hours and sfTtime is less than 21.96 hours, the possibilities of owners' decisions to share or not are quite close, with values of 53.9% for not sharing and 46.1% for share.

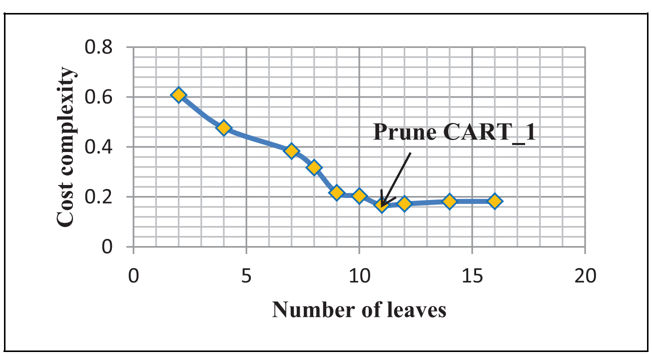


**Figure 3.** Procedure of selecting an optimal CART_1.

Likewise, Node 3 can be further split by the narrowed intervals of sfDtime and sfTtime, forming Node 6, LeafNode 4, Node 7, LeafNode 9, and LeafNode 10. The variety of subclasses also reflects the complexity of the decision-making process affected by owners' self-use parking time schedules. Afterwards, Node 6 directs types of less than 0.5, namely, underground spots to the left and on-street spots to the right, forming LeafNode 2 and LeafNode 3, respectively. The predicted results reveal that owners of on-street spots are more likely to share than those of underground spots. In Node 7, values of sfDfres lower than 2.02 are to the left, forming Node 8, and those above 2.02 to the right, forming LeafNode 8. The predicted result demonstrates that owners who park more than 2.02 times during the daytime on workdays are more prone to share than those who might be parking for a long period without moving. CART_1 further proceeds to split Node 8 by Floor, where values of Floor less than 2.50 form LeafNode 5, which indicate that the spots on the street and in B1 are inclined to be unshared under that specific circumstance. However, the attitudes are not clear for owners in B2, as is reflected in Node 9 (53.3% versus 46.7%). Thus, Node 9 is further split by sfDtime at the value of 8.22 hours. Values of sfDtime less than 8.22 hours tend to be unshared, whereas those exceeding 8.22 hours are more likely to share, because their parking time primarily focuses on the daytime hours, thereby leaving the nighttime available.

Lastly, a set of sharing decision rules, which can be identified by computers and adopted by managements, were extracted from 11 nodes of CART_1. The detailed rules are listed in Table 2, and it can be observed that six rules refer to share and five rules refer to unshare. The total predicting accuracy of CART_1 is 93.4%; furthermore, CART_1 performs better in predicting unshare, rather than share (97.8% versus 89.0%). To validate the results, the same dataset is adopted in a widely accepted statistical model, the binary logistic model; and the prediction accuracy is lower than that of CART_1 with similar findings.

## *CART_2: How Long to Share during the Peak of Parking Demand?*

To predict the sharing time during the peak period of parking demand, namely, in the daytime on workdays, 377 shared samples were utilized to build CART_2. Apart from the independent variables employed in CART_1, the rental effects of the previous month's variables were also adopted in this model. Moreover, the dependent variable here is the continuous variable osDtime.

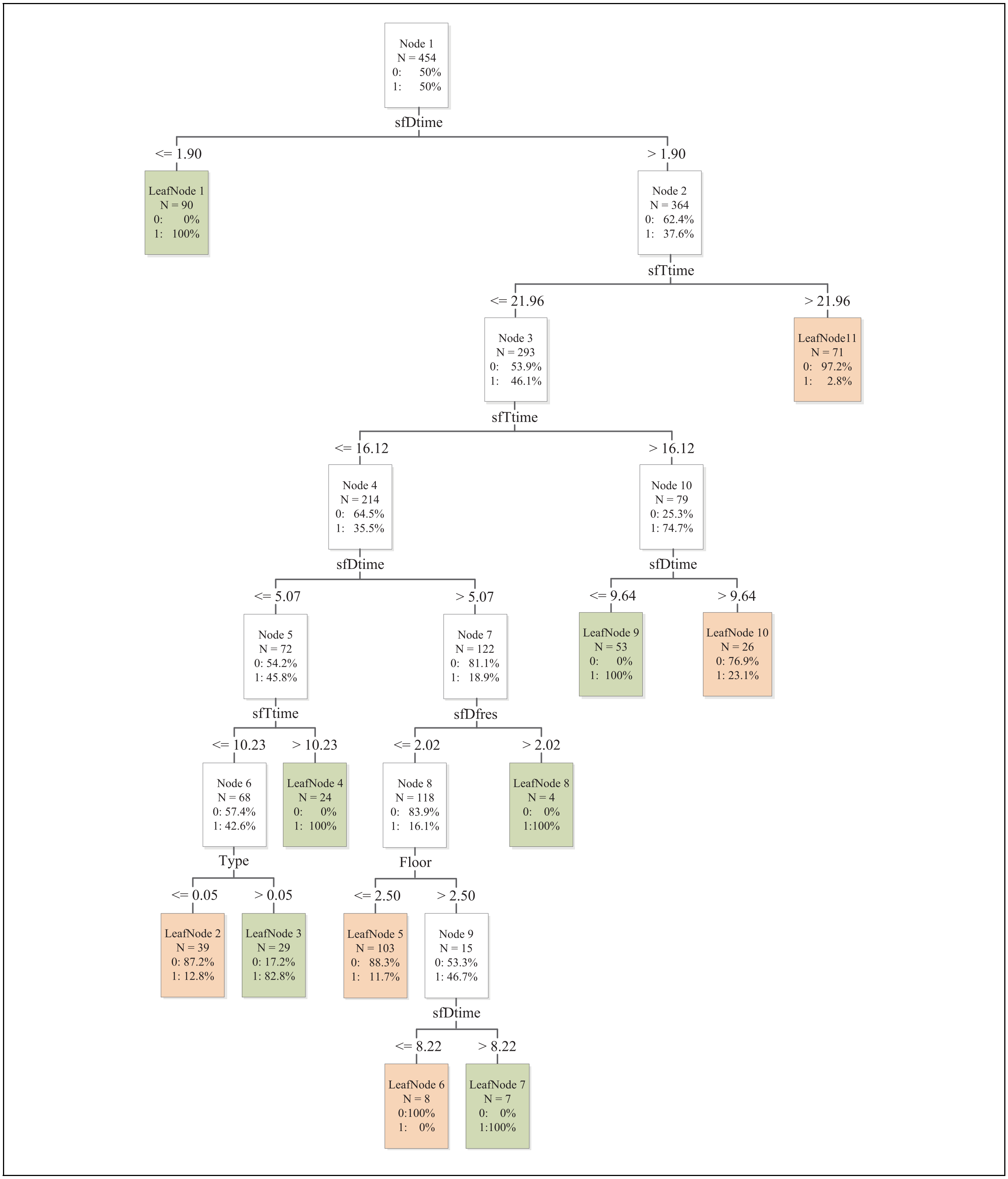


**Figure 4.** Graphic expression of CART_1.

**Table 2.** Production rules in CART_1

| No. | Antecedent | Predicted osYN | Accuracy |
|---|---|---|---|
| 1 | sfDtime<=1.90; | share | 100% |
| 2 | 1.90<sfDtime<=5.07; sfTtime<=10.23; Type<=0.05; | unshare | 87.2% |
| 3 | 1.90<sfDtime<=5.07; sfTtime<=10.23; 0.05<Type; | share | 82.8% |
| 4 | 1.90<sfDtime<=5.07; 10.23<sfTtime<=16.12; | share | 100% |
| 5 | 5.07<sfDtime; sfTtime<=16.12; sfDfres<=2.02; Floor=<2.50; | unshare | 88.3% |
| 6 | 5.07<sfDtime<=8.22; sfTtime<=16.12; sfDfres<=2.02; 2.50<Floor; | unshare | 100% |
| 7 | 8.22<sfDtime; sfTtime<=16.12; sfDfres<=2.02; 2.50<Floor; | share | 100% |
| 8 | 5.07<sfDtime; sfTtime<=16.12; 2.02<sfDfres; | share | 100% |
| 9 | 1.90<sfDtime<=9.64; 16.12<sfTtime<=23.18; | share | 100% |
| 10 | 9.64<sfDtime; 16.12<sfTtime<=21.96; | unshare | 76.9% |
| 11 | 1.90<sfDtime; 21.96<sfTtime; | unshare | 97.2% |

*Note*: No. = number.

Based on the same preset parameters (maximum leaf numbers and maximum tree depth) and testing methodology as in CART_1 as well as the impurity reduction thresholds for RT (illustrated in Method), a pruned CART_2 that leads to the lowest cost complexity for the testing data was determined. Figure 5 depicts the relationship between the number of leaves and cost complexity in CART_2. The cost complexity reached the bottom at 15 leaves, with a value of 0.296.

CART_2 has 15 leaf nodes, where the shaded rectangles represent different levels of a willingness to share. As the total daytime period consists of 12 hours, those between 8 and 12 hours are classified as high-level in the red rectangles; those between 4 and 8 hours are classified as mid-level in the blue rectangles, and those below 4 hours are classified as being at a low-level in the green rectangles.

As graphically presented in Figure 6, CART_2 initially split at Node 1(the root node), based on sfDtime, forming Node 2 with an average sharing time of 8.99 hours and Node 9 of 2.49 hours. The significant difference in sharing time demonstrates that self-use parking time represents a priority when owners consider their sharing willingness during peak periods of parking demand. Node 2, with values of sfDtime less than 1.82 hours, is further split by sfDfres. Values of sfDfres less than 1.57 are gathered in the left child Node 3, whereas those above 1.57 are in the right child Node 7. The difference in the average sharing time in Node 3 and Node 7 (9.58 hours versus 7.86 hours) indicates that self-use parking frequency might be beneficial, so as to distinguish the high-level from the mid-level sharing time. Afterwards, Node 3 and Node 7 are further split by the rental effect of the previous month's variables rtTtime and rtosDR, as well as spots' physical characteristics variables distanceA and distanceL, forming LeafNodes 1–8. Likewise, CART_2 further split Node 9 based on the variables rtDtime, sfDtime, rtosDR, distanceL, and sfTtime into LeafNodes 9–15, shown on the right-hand side of the tree in Figure 6. Compared with LeafNodes 1-8 on the left-hand side, values of sfDtime above 1.82 exhibit a lower sharing willingness. It is also shown that all five high-level leaf nodes are on the left, whereas all four low-level leaf nodes are on the right.

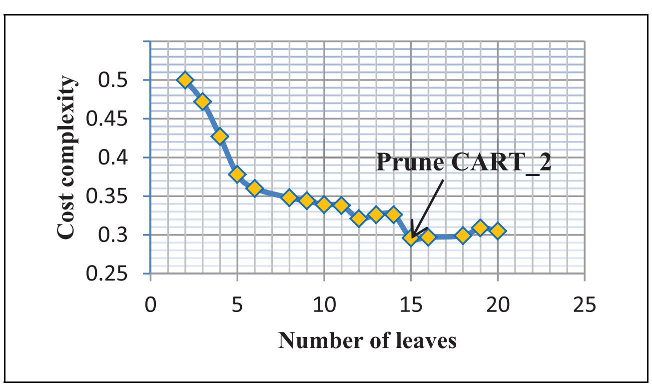


**Figure 5.** Procedure of selecting an optimal CART_2.

Judging from the split values and the average sharing time of the child trees, it can be inferred that a higher rental time in the previous month would boost owners' willingness to share this month, such as in Node 3, where the average sharing time is 11.11 hours for those values of rtTime more than 5.55 hours, as compared with 8.78 hours for those values of rtTime less than 5.55 hours. The same situation can be found in Node 7, Node 5, and Node 9. However, rtosDR is not an important consideration for owners, as the lower rtosDR of the previous month is related to a higher sharing willingness this month in Node 6 and Node 11. Moreover, the lower self-use parking time leads to a higher sharing time, as indicated in Node 1, Node 10, Node 12, and Node14. Another inference is that closer distances to surrounding buildings would enhance the owners' willingness to share, as in Node 4, Node 8, and Node 13.

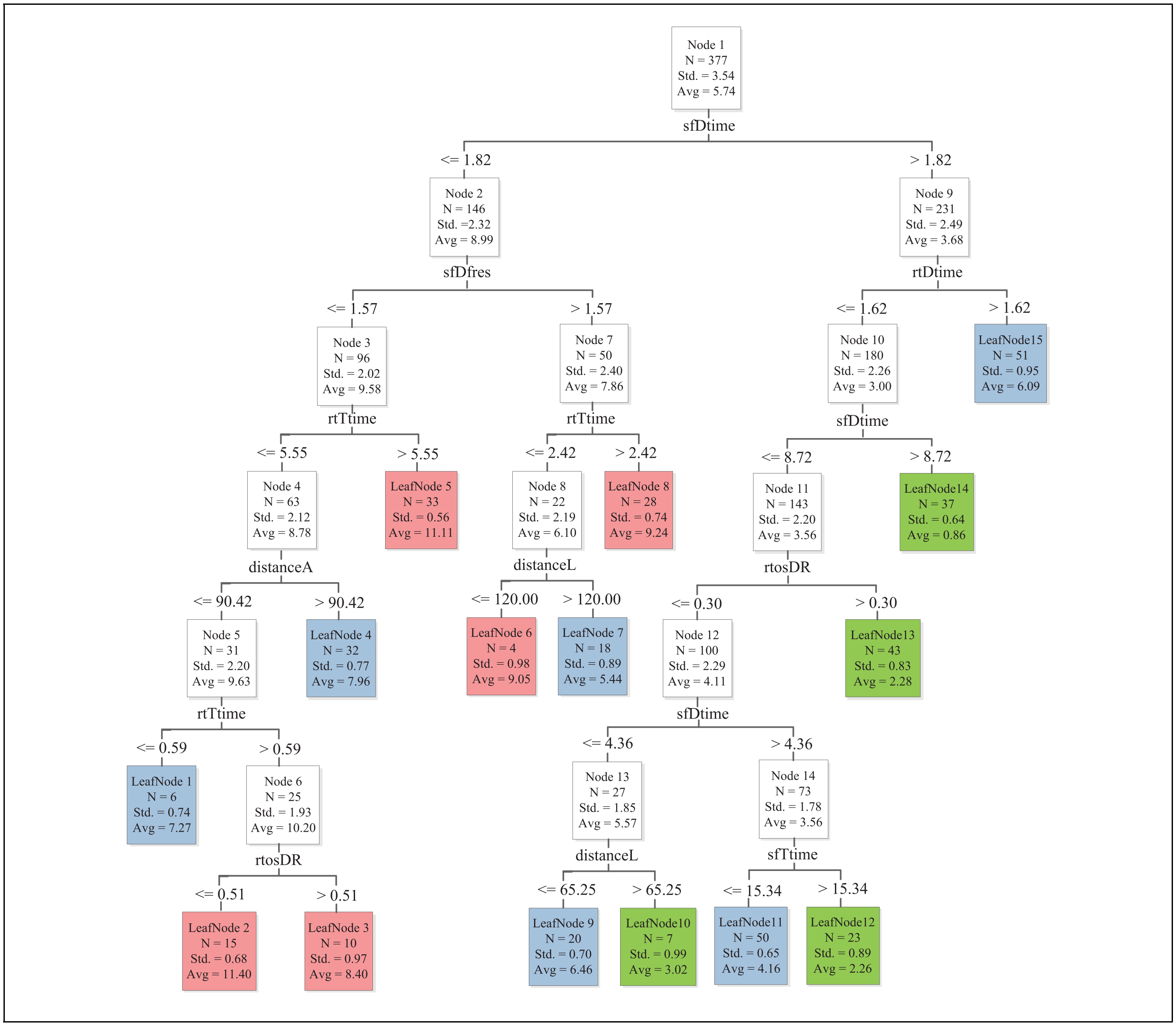


**Figure 6.** Graphic expression of CART_2.

Lastly, a set of sharing decision rules, which can be readily identified by computers and adopted by managements, was extracted from 15 nodes of CART_2, and the detailed rules are revealed in Table 3. It can be observed that five rules refer to a high-level sharing willingness, six rules to a mid-level and four rules to a low-level. The total predicting *R*-squared of CART_2 is 0.95. Therefore, CART_2 performs significantly better in high-level sharing time (*R*-squared = 0.96) than in a mid-level (*R*-squared = 0.95) or low-level (*R*-squared = 0.94). The same dataset is also adopted in a traditional statistical model, the linear regression model; and the *R*-squared is lower than that of CART_2 with no new findings of variables.

## Discussions on Predictor Variables

Although the research purposes of CART_1 and CART_2 are quite diverse, considering the input variables are somewhat similar, the importance of variables (seen in Table 4) in the two models can be compared, along with several findings that are illustrated as follows.

All three variables of owners' self-use behavior, including sfDtime, sfTtime, and sfDfres, are mutually significant factors in both CART_1 and CART_2. Most especially, sfDtime wins a score of 100.00 in both models, meaning the self-use time during the peak periods of parking demand is certainly the most vital consideration in both decision-making processes. Judging from the order and importance of the variables, sfTtime and

**Table 3.** Production Rules in CART_2

| No. | Antecedent | Predicted osDtime | SD |
|---|---|---|---|
| 1 | sfDtime<=1.82; sfDfres<=1.57; rtTtime<=0.59; distanceA<=90.42; | 7.27 | 0.74 |
| 2 | sfDtime<=1.82; sfDfres<=1.57; 0.59<rtTtime<=5.55; distanceA<=90.42; rtosDR<=0.51; | 11.40 | 0.68 |
| 3 | sfDtime<=1.82; sfDfres<=1.57; 0.59<rtTtime<=5.55; distanceA<=90.42; 0.51<rtosDR; | 8.40 | 0.97 |
| 4 | sfDtime<=1.82; sfDfres<=1.57; rtTtime<=5.55; 90.42<distanceA; | 7.96 | 0.77 |
| 5 | sfDtime<=1.82; sfDfres<=1.57; 5.55<rtTtime; | 11.11 | 0.56 |
| 6 | sfDtime<=1.82; 1.57<sfDfres; rtTtime<=2.42; distanceL<=120.00; | 9.05 | 0.98 |
| 7 | sfDtime<=1.82; 1.57<sfDfres; rtTtime<=2.42; 120.00<distanceL; | 5.44 | 0.89 |
| 8 | sfDtime<=1.82; 1.57<sfDfres; 2.42<rtTtime; | 9.24 | 0.74 |
| 9 | 1.82<sfDtime<=4.36; rtDtime<=1.62; rtosDR<=0.39; distanceL<=65.25; | 6.46 | 0.70 |
| 10 | 1.82<sfDtime<=4.36; rtDtime<=1.62; rtosDR<=0.39; 65.25<distanceL; | 3.02 | 0.99 |
| 11 | 4.36<sfDtime<=8.72; rtDtime<=1.62; rtosDR<=0.39; sfTtime<=15.34; | 4.16 | 0.65 |
| 12 | 4.36<sfDtime<=8.72; rtDtime<=1.62; rtosDR<=0.39; 15.34<sfTtime; | 2.26 | 0.89 |
| 13 | 1.82<sfDtime<=8.72; rtDtime<=1.62; 0.39<rtosDR; | 2.28 | 0.83 |
| 14 | 8.72<sfDtime; rtDtime<=1.62; | 0.86 | 0.64 |
| 15 | 1.82<sfDtime; 1.62<rtDtime; | 6.09 | 0.95 |

*Note*: No. = number; SD = standard deviation.

**Table 4.** Importance of Variables in Two Models

| CART_1 | | CART_2 | |
|---|---|---|---|
| Variables | Score | Variables | Score |
| sfDtime | 100.00 | sfDtime | 100.00 |
| sfTtime | 63.19 | rtDtime | 59.32 |
| sfDfres | 18.22 | rtTtime | 21.55 |
| Floor | 13.45 | rtosDR | 18.70 |
| Type | 8.21 | distanceL | 16.43 |
| | | sfDfres | 11.21 |
| | | sfTtime | 10.94 |
| | | distanceA | 5.73 |

sfDfres are more important in CART_1 than in CART_2. It indicates that owners may evaluate their openness-to-share decision, based primarily on their own individual parking behavior, while paying less attention to self-use parking, once they have made their decisions to share.

Moreover, the influences of the self-use variables in the two models are not quite consistent. Theoretically, the increase of self-use time would lower owners' sharing willingness. This was proved true in CART_2 (see Node 1 and Node 12). However, in CART_1, the intervals of sfDtime were split several times and the open-to-share possibility was simply not consistent with sfDtime. It indicates that the decision-making process of whether to share or not is rather complex, and the owners' willingness to share most probably resulted from the available variety of self-use parking modes.

As for spots' physical characteristics, Floor and Type are testified important to CART_1 and the other two variables, distanceL and distanceA, are essential to CART_2 (see Table 4). The disparity indicates that owners most likely value the basic construction characteristics more when considering whether to share or not, whereas in sharing practices, they may determine that the location of the spots is more important to rental effects. Therefore, their sharing willingness might be influenced more significantly by the two distance variables.

The most profound difference is the participation of variables in the rental effects of the previous month group in CART_2. All three variables ranked from second to fourth on the importance of variables list, which reveals that owners' willingness to share may be highly encouraged by the rental effects of the previous month.

## Conclusions

Shared parking of private residential parking spots is an innovative policy in China. Furthermore, it can substantially improve the utilization of parking resources and profoundly relieve the pressure on parking supply, especially during the peak periods of parking demand. Based on the real data of 1-year behavioral records of owners obtained from applications on smart phones, as well as field survey data, the study presented in this paper analyzed the influential factors in owners' sharing decision mechanism. Moreover, two CART models were developed to adequately predict the willingness of spot owners to share private parking. CART_1 answers whether owners would share their spots, whereas CART_2 explains the average time owners would be willing to share during the peak periods of parking demand.

The major conclusion is that owners' self-use behavior, spots' physical characteristics and rental effects of the previous month have all been found to have a significant influence on owners' willingness to share. More specifically, it was determined that the owners always put more emphasis on their own parking demands ahead of sharing decisions. Moreover, the self-use parking modes significantly affect the owners' start-to-share decision. Once shared, the owners' sharing willingness during the peak periods of parking demand would largely rely on the previous month's rental effects, if the self-use parking demand is satisfied. Furthermore, the influential factors of spots' physical characteristics vary according to the two models, thereby indicating that the consideration is dissimilar in the two stages.

It probably is the first time that a spot-level study of owners' sharing willingness driven by real-data and modeled by both CT and RT has been conducted. The main contribution of CART_1 is to facilitate the classification of owners into several groups, based on their self-use parking habits and spots' physical factors as well as to assess the openness-to-share possibility of each group. The main contribution of CART_2 is to adequately identify the important factors affecting owners' sharing willingness during the peak periods of parking demand and, furthermore, to predict the sharing time of each owner. A combination of the two models can effectively predict the total potential of shared parking in a residential zone, which would substantially benefit the government, as relating to their parking supply policies, and also be of advantage to third parties to enhance the distribution of parking resources.

The study has the following limitations. First, the owners' willingness to share may vary according to human factors, such as age and family composition, which this study did not expound on, because of confidentiality agreements aimed at protecting the owners' privacy. Future studies will optimize the data collection methodology and acquire more demographic information by means of door-to-door canvassing in other study regions. Second, the owners' willingness to share a parking spot was not correlated with those in other residential communities; therefore, in future studies, more communities with diverse geographic land use and demographic attributes should be researched and systematically compared and evaluated.

## Acknowledgments

This research is supported by the National Natural Science Foundation of China (51478111, 51508094). The authors would like to express their appreciation to the Beijing Tongyudao Technology Co., Ltd. for their data support.

## Author Contributions

Study conception and design: Jun Chen, Chu Zhang; data collection: Chu Zhang, Yuanyuan Wu; analysis and interpretation of results: Zhibin Li, Chu Zhang; draft manuscript preparation: Chu Zhang.

All authors reviewed the results and approved the final version of the manuscript.

## References

1. Weinberger, R., M. Kodransky, J. Karlin-Resnick, A. Gauthier, Z. Gyarmati, Y. Pen, and L. Fu. *Parking Guidebook for Chinese Cities*, Handbooks, 2013.
2. Tian, G., R. Ewing, R. Weinberger, K. Shively, P. Stinger, and S. Hamidi. Trip and Parking Generation at Transit-Oriented Developments: A Case Study of Redmond Tod, Seattle Region. *Transportation*, Vol. 44, 2017, pp. 1235–1254.
3. Chaniotakis, E, and A. J. Pel. Drivers' Parking Location Choice under Uncertain Parking Availability and Search Times: A Stated Preference Experiment. *Transportation Research Part A: Policy and Practice*, Vol. 82, 2015, pp. 228–239.
4. Ibeas, A., L. Dell'Olio, M. Bordagaray, and J. D. Ortúzar. Modelling Parking Choices Considering User Heterogeneity. *Transportation Research Part A: Policy and Practice*, Vol. 70, 2014, pp. 41–49.
5. Li, Z. B., W. Wang, C. Yang, and D. R. Ragland. Bicycle Commuting Market Analysis Using Attitudinal Market Segmentation Approach. *Transportation Research Part A: Policy and Practice*, Vol. 47, 2013, pp. 56–68.
6. Efthymiou, D., C. Antoniou, and P. Waddell. Factors Affecting the Adoption of Vehicle Sharing Systems by Young Drivers. *Transport Policy*, Vol. 29, 2013, pp. 64–73.
7. Shaaban, K., and A. Pande. Classification Tree Analysis of Factors Affecting Parking Choices in Qatar. *Case Studies on Transport Policy*, Vol. 4, No. 2, 2015, pp. 88–95.
8. Rogers, J., D. Emerine, P. Haas, D. Jackson, P. Kauffmann, R. Rybeck, and R. Westrom. Estimating Parking Utilization in Multifamily Residential Buildings in Washington, D.C. *Transportation Research Record: Journal of the Transportation Research Board*, 2016. 2568: 72–82.
9. Xu, S. X., M. Cheng, X. T. Kong, H. Yang H, and G. Q. Huang GQ. Private Parking Slot Sharing. *Transportation*

*Research Part B: Methodological*, Vol. 93, 2016, pp. 596–617.
10. Shao, C., H. Yang, Y. Zhang, and J. Ke. A Simple Reservation and Allocation Model of Shared Parking Lots. *Transportation Research Part C: Emerging Technologies*, Vol. 71, 2016, pp. 303–312.
11. Smith, M. S. *Shared Parking*. 2nd ed. Urban Land Institute, Washington, D.C., 2005.
12. Meng, Q., and J. Weng. Classification and Regression Tree Approach for Predicting Drivers' Merging Behavior in Short-Term Work Zone Merging Areas. *Journal of Transportation Engineering*, Vol. 138, 2012, pp. 1062–1070.
13. Chang, L. Y., and H. W. Wang. Analysis of Traffic Injury Severity: An Application of Non-Parametric Classification Tree Techniques. *Accident Analysis and Prevention*, Vol. 38 2006, pp. 1019–1027.
14. Breiman, L., J. H. Friedman, R. A. Olshen, and C. J. Stone. *Classification and Regression Trees*. Wadsworth Brooks/Cole, Pacific Grove, CA, 1984.
15. Weng, J., S. Xue, and X. Yan. Modeling Vehicle Merging Behavior in Work Zone Merging Areas During the Merging Implementation Period. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 17, 2016, pp. 917–925.
16. Kashani, A. T., and A. S. Mohaymany. Analysis of the Traffic Injury Severity on Two-Lane, Two-Way Rural Roads Based on Classification Tree Models. *Safety Science*, Vol. 49, 2011, pp. 1314–1320.
17. Lewis, D. D., and J. Catlett. Heterogeneous Uncertainty Sampling for Supervised Learning. *Machine Learning Proceedings*, 1994, pp. 148–156.

*The Standing Committee on Transportation Demand Management (ABE50) peer-reviewed this paper (18-02688).*