

An Application of Intelligent Noise Filtering Techniques in Demand Forecasting for Carsharing Systems[#]

Hooi Ling KHOO^{*1} Chau Ha FUNG^{*2&} Qiang MENG^{*3} Der-Horng LEE^{*4}

*Department of Civil Engineering^{*1}*

National University of Singapore

(Block E1, #08-20, 1 Engineering Drive 2, Singapore 117576.

TEL: (+65) 6516- 5035, E-mail: hooiling@nus.edu.sg)

*Credit Suisse^{*2&}*

(Block 77, #06-523, Commonwealth Drive, Singapore 140077.

TEL: (+65) 93389485, E-mail: jenicefung@gmail.com)

*Department of Civil Engineering^{*3}*

National University of Singapore

(Block E1, #07-10, 1 Engineering Drive 2, Singapore 117576.

TEL: (+65) 6516- 5494, E-mail: cvemq@nus.edu.sg)

*Department of Civil Engineering^{*4}*

National University of Singapore

(Block E1A, #07-16, 1 Engineering Drive 2, Singapore 117576.

TEL: (+65) 6516- 2131, E-mail: cveleedh@nus.edu.sg)

This study deals with demand forecasting for the carsharing system with multi-station and flexible returning station and time. Neural network is employed as the simulation model to forecast the demand at each station at certain time period in the system. This study introduces intelligent filtering techniques as a tool to remove the noise of the data before it is fed into the simulation model. Two filtering techniques have been tested, namely outlier analysis and cluster analysis. Results show that outlier analysis is better compared to cluster analysis in enhancing the accuracy of forecasting model. This shows that proper choice of techniques is important to guarantee that the introduction of this extra procedure could improve the forecasting accuracy.

Keywords: *Simulation, Data Mining, Demand forecasting*

1. Introduction

Traffic congestion is a worldwide problem that has been plaguing traffic researchers, engineers and authorities for a few decades. Various advanced technologies such as GPS, sensors and communication technologies have been adopted to help better control of traffic on the roads. However, if the vehicle ownership and usage is growing faster than what the traffic system can handle, these control policies are destined to fail at some point of the time. As a complement to a more efficient traffic management, there is a need to advocate the usage of public transportation system in order to control the vehicle ownership. Nevertheless, public transportation system is not attractive to the travelers. It is because it cannot provide the comparable convenience, comfort and privacy to the travelers as private vehicles do. Hence, another means of transport which can capture

the benefits of these two, that is, it does not increase the vehicle ownership while at the same time, it carries all the features that a private vehicle can provide, is much anticipated. This trend has stimulated the emergence of carsharing system. With this system, authorities are able to control the growing of the vehicle population and in the meanwhile, the commuters are able to enjoy their trip with hassle-free. It is a win-win strategy. Currently, there are several of these activities being taken place in the big cities in the world such that about 40 programs were deployed in North America (1), 18 in Japan and 4 in Singapore (2).

Carsharing system is referred as a system with large number of people to share a small number of vehicles that are reserved for them and used individually as required (3). Users pay only for the time used and distance traveled, while the carsharing company owns the vehicles and handle all the repairs and insurance. This should be distinguished from carpooling and car

[#] This paper won the second prize of the Best Scientific Paper in the 14th World Congress on ITS, Beijing, 2007

&The work is carried out when she was a Research Engineer in Department of Civil Engineering, National University of Singapore

rental where in the former, vehicles are shared for some time but the main ownership and arrangements remain unchanged, and in the latter, vehicles are rented over an extended period of time (4). According to some recent research studies, carsharing system is found to be able to improve travelers' mobility, lowering emissions, reducing congestion problems (5) and reducing vehicle ownership in which each carsharing vehicle replaces 14.9 privately owned vehicles (6). Some of the well known carsharing systems are such as IntelliShare system in California (7), (8), Praxitele in Europe (9) and Honda Diracc in Singapore (10). In fact, carsharing system is not a very new policy; its application can be traced back to 1948 where the first carsharing project was launched in Europe. A more detail elaboration of the history of carsharing system and its benefit can be found in (4).

1.1. Overview of Honda Diracc

Established in March 2002, Honda Diracc has now emerged as an important operator of car sharing system in Singapore (10). Over the years, Honda Diracc has experienced continuous growth in terms of number of stations, fleet size and membership. Currently, there are 19 stations, where most of them are located at the central business district (CBD) (10). It operates using environment-friendly Honda Civic Hybrid vehicles, with over 50 vehicles in the scheme. More than 2000 members have enrolled in the program, with over 70 trips being made in an average day. The system adopted advanced technologies such as GPS and RFID vehicle location systems, contactless smartcard access, internet or phone-based reservations and touchscreen display units. Communicating with the backend computer system periodically, information from these vehicles prompts the system manager to relocate vehicles between stations when needed. The relocation algorithm periodically checks the number of available vehicles at each station against a set of predefined threshold values and triggers a relocation trip if the number of available vehicles in certain station hits the lower or upper boundary.

The key difference that makes Honda Diracc distinguished from the other three competitors in Singapore is that it offers the flexibility of picking-up and returning vehicles at any station based on users' choice. The returning station specified by a user prior to a trip may be changed en-route. They are also not required to state their return times. This implies that a user can drive a car from station A and return it to station B after used. He can rent the car for as short as 10 minutes or as long as a few days. With this freedom offered to the users, the operators are constantly challenged to maintain an even distribution of vehicles in all stations with an efficient relocation system. A delicate balance needs to be struck between having

available vehicles for pick-up and having enough empty space for other vehicles to return to.

1.2. Intelligent noise filtering techniques

To facilitate the relocation arises from the flexibility of the system, the demand and the supply of each of the station needs to be forecasted over time. The demand of vehicles is defined as the number of vehicles being rented from a station while the supply of vehicles is defined as the number of vehicles being returned to the station during a particular time period. By predicting the demand and supply, we may know that whether there is excessive demand or supply at a station. Excessive demand indicates that the station need more vehicles while excessive supply means that some vehicles in the station need to be driven away as the parking spaces rented by the operator at the station is fixed. Both conditions initiate the need for relocation. Since the operators have no control over the demand and the supply of the number of vehicles in a station, there is a need for them to forecast the demand and supply in order to ascertain the efficiency of the system. Hence, forecasting is a key issue in this flexible system. The scope of this study however is limited to demand forecasting only although the techniques presented herein can also be applied to supply and vehicle net flow forecasting.

There is very little literature study can be found on the forecasting of carsharing demand. In year 2001, Barth and Todd (7) have conducted a research study on the user behaviour in the IntelliShare carsharing system, where the return time of vehicles is known. In their study, a Chi-square goodness-of-fit test was applied to statistically compare the trip frequency patterns of a single day with the running average. Nonetheless, this study is different from the case of Honda Diracc (in this study) because we do not have the information about the return time of vehicle since the operator gives the freedom of using the vehicle for as long as they want without having to specify a time of return in advance. Based on data type and horizon of forecasts, Kek (4) has selected the linear models, i.e. the Holt's method and Selective Moving Averages (SMA) as the demand forecasting tool. She found out that Holt models consistently outperforms SMA model. She adopted non-linear model, i.e. the neural network model as well. In the later research, she adopted multiple regression model to do the forecasting (11). Cheu et al. (12) had conducted a research study on forecasting the vehicle net flow in carsharing system using two approaches, namely the neural networks and the support vector machine. In their study, a comparison of the performance of these models is carried out. By using six months of data, it is found that the multilayer perceptron neural network model slightly outperformed the support vector machine and has better accuracy in forecasting.

One pitfall of these studies is the data is not processed prior to the adoption. The preprocessing data treatment is vital during the forecasting due to two main reasons. Firstly, this is done to remove noise inside the data set. The data obtained for demand forecasting in carsharing system is much depended on human behavior. Different individual (member) will exhibit different usage habit and they cannot be represented by a general relationship. There exist some people that may not comply with the general behavior of the rest of the members' data. Inclusion of such data set in the forecasting model may lead to misleading results. Secondly, by doing this, the forecasting model can be enhanced and the accuracy can be improved. Hence, prior to the simulation, these data needed to be filtered in order to remove the noise component. In this study, two filtering techniques that are adopted are outlier analysis and clustering analysis while the forecasting model used is neural networks.

Outlier analysis technique (13) deals with the sample data that does not comply with the general behavior of the data model, which is termed as outliers. In this study, outlier data is caused by the difference in member behavior in carsharing. Outlier analysis is about removing the outlier data from a data set and can be described as a process of the selection of k out of n samples that are considerably dissimilar, exceptional, and inconsistent with respect to the remaining data. There are various methods in detection of outliers data. To name a few, they are data visualization methods, statistic method, distance-based outlier detection and deviation-based method. The choice of the method depends on the number of dimension of the data. For example, statistic method is appropriate for 1-dimensional data but not multidimensional data.

Cluster analysis is a set of methodologies for automatic classification of samples into a number of groups using a measure of association, so that the samples in one group are similar and samples belonging to different groups are not similar. The input for cluster analysis is a set of samples and a measure of similarity between two samples. The output from cluster analysis is a number of groups that form a partition, or a structure of partitions, of the data set (13). This method is suitable for the exploration of interrelationship among samples to make a preliminary assessment of the sample structure. It is important to mention that there is no clustering technique that is universally applicable in uncovering the variety of structures present in data sets. There are two popular clustering algorithms, namely the hierarchical clustering and the iterative square-error partitional clustering. The choice of the appropriate method however is much depended on the users' understanding of the problem and the corresponding data types.

1.3. Objectives of the study

The objective of this study is to evaluate the

performance of the forecasting model when the filtering techniques are applied to remove noise data prior to simulation. The forecasting simulation model adopted in this study is the well known neural networks. Secondly, it is aimed to compare the effectiveness of different type of filtering techniques specifically to the demand forecasting in the carsharing system. These filtering techniques are outlier analysis and cluster analysis. To the authors' best knowledge, this paper is the first documented attempt to adopt these preprocessing techniques to treat data in demand forecasting of carsharing system with flexible return station and time.

2. Methodology

2.1. Demand forecasting framework

The demand forecasting simulation model adopted in this study is neural network. The type of model used in simulation is Multi-Layer Perceptron (MLP) Neural Network. It consists of a single input layer where the input variables are being entered, a single hidden layer with a variable number of hidden neurons and a single output layer. The network architectures applied is feedforward networks. It consists of one or more hidden layers of sigmoid neurons followed by an output layer of linear neurons. Multiple layers of neurons with nonlinear transfer functions allow the network to learn nonlinear and linear relationships between input and output vectors. The training type adopted is batch training which means that the weights and biases are being updated only after all of the inputs are presented. The procedure of the demand forecasting model is explained as follows:

- Step 1: (Initialize) Define three simulation data sets: training data set, validation data set and test data set.
- Step 2: (Filtering) Filter all the data in Step 1 using the associated intelligent filtering techniques. Refer to Figure 1.
- Step 3: (Update Neural Network) The data from training set is used to train the model so that the weightings in neurons are being updated from time to time.
- Step 4: (Validation) Use the trained network to simulate the validation data set and compare with the corresponding output.
- Step 5: (Optimality Checking) Check whether the mean square error of the validation set has reached minimum? If yes, go to Step 6 and retain the value of weights and biases parameters; otherwise go to Step 3.
- Step 6: (Input test set data) Simulation of test data set is carried out using the weights and biases values obtained from Step 5.
- Step 7: (Termination) End the simulation and collect the forecast result.

There are three models developed in this study. The above mentioned procedure is the general steps to be followed by all models except Step 2. At Step 2, different model will have different filtering techniques as well as filtering data set. For example, `nn_Outliers` only has its training data set being filtered using the outlier analysis but not for validation set and test set data. For the `nn_Cluster` model, all input data sets are filtered prior to simulation. For `nn` model which serves as the benchmark model, has the all-pass filter. All-pass filter is equivalent to no filtering as this filter allows all the data to pass through. This means that the input and output after filtering is the same. Figure 1 illustrates these models.

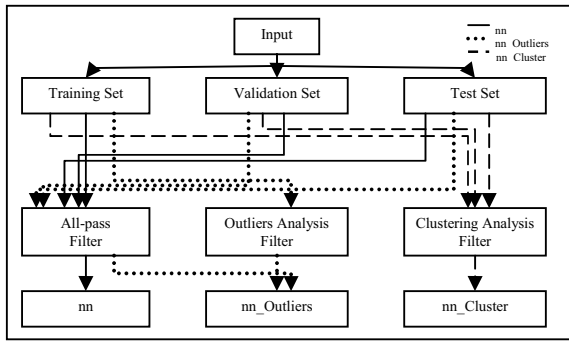


Figure 1: Input data filtering process

2.2. Simulation set-up

2.2.1. `nn` Model. For this all-pass filter model, the filtered output data will be the same as input data as there is no cleaning process undergoing inside. The inputs of the model are the factors having strong correlations with trip frequency and they were used as input variables in the model (4). Essentially, they are:

- i. Demand at the same station during the same time period, on the same day of the previous week.
- ii. Demand at the same station during the same time period on the previous day
- iii. Demand at the same station during the previous time period
- iv. Demand at the same station during the second previous time period.

2.2.2. `nn_Outliers` Model. It is observed that members' trip patterns are strongly related to their habitual behaviors. Some members are frequent users in the program while others only make a trip once in a while. Since it is hard to trace the infrequent usage members' habitual usage pattern, the trips they made may be considered as outliers and is removed in the training set. Therefore, the outliers filter removes the trips (data) made by infrequent users in the carsharing system while only the trips made by frequent users is used for training purposes. This is because the trips made by the frequent

users tend to be regular and the simulation model can be trained by this more specific data set that is less affected by noise. Using the data visualization method, the frequent users are defined as the members that fulfill either one of the following criteria:

- i. Those members who have made more than 20 trips if they join as member before 2006
- ii. Those members who have made more than 5 trips if they join as member in 2006.

Those who do not fulfill either of these criteria are treated as outlier and their associated data are removed from analysis. The input variables of this model are the same as the input of `nn` model except that the outliers data is removed.

2.2.3. `nn_Cluster` Model. The main task of this model is to find out the relationships among the different periods of a station using the clustering analysis. For the benchmark model, all the stations are assumed to have similar usage pattern over all time period. Hence, the input of the simulation model for all the stations and time periods is identical. From the experience, we know that the demand of vehicles in each period is different for different stations. The activity of each station is highly depended on its location, its neighbourhood buildings and the road traffic conditions vicinity of the stations. Hence, it is more rational to consider individual time period of each station differently so that it gives a more comprehensive analysis on how relationships between periods vary among different stations and different trip flow data cases. The application of cluster analysis as filtering technique aims to find the specific input factors for individual station. In this study, K-means clustering method is adopted and the procedure is such as follows:

- Step 1: (Initialize) Randomly assign each period of data to a cluster and set the initial position for the cluster centroid.
- Step 2: (Calculate Euclidean distance) The Euclidean distance for all period data in the cluster is calculated using equation (1) and assign the distance to each period data.

$$d = |\mathbf{x} - \mathbf{y}| = \sqrt{\sum_{i=1}^N |x_i - y_i|^2} \quad (1)$$

where x_i is the coordinate for the i^{th} centroid,

y_i is the i^{th} sample and d is the Euclidean distance.

- Step 3: (Assign period to nearest centroid) Each period data is assigned to the nearest centroid. When the samples are assigned into a new cluster, a new centroid is obtained by taking the mean of the centroid coordinate.
- Step 4: (Error checking) The error (the distance from period data to centroid) is calculated and is

compared to the pre-specified error. If the error is less than the pre-specified one, stop; else go to Step 2.

A row of input data from previous period to two periods before the previous week is used for clustering. The duration of data used for clustering is from July 2005 to March 2006 so that the forecasting for April 2006 could be done. The number of clusters is set from three clusters to nine clusters and the procedure mentioned beforehand is carried out. At the end of the day, it is expected that different period of time will have different period of input variables.

2.3. Data description

The data set used for analysis in this study is obtained from Honda Diracc in Singapore. By April 2006, there are a total of 12 stations with the fleet size of 60 vehicles is operating in Singapore. The majority of stations are densely located in Central Business District (CBD), with only one port located at airport. There are about 2000 members participated in this program. The task here is to forecast the demand for April, 2006 using historical data. In order to train the neural network model, the data from July, 2005 to February, 2006 excluding December and January data, is used as the training set data while data for March, 2006 is used for model validation purposes. The month of December and January are not included because of its unstable features due to the festive seasons, which are observed during the preliminary analysis. Forecast is made for each time period in April, in such a way that each time period consists of 2 hours in a rolling manner. For instance, the 2-hourly rolling time periods for period 1 stands for the time from 0100 to 0300, period 2 stands for time from 0200 to 0400 and so on. Due to the low demand of vehicles in early morning and late evening time periods, forecast are done from period 7 to period 22 only (whole day data). The time periods are further categorized into morning: period 7 to period 11, afternoon: period 12 to period 17 and evening: period 18 to period 22 so that comparison can be made according to different time periods.

3. Results and Discussions

3.1. Measures of effectiveness (MOE)

The MOE used to evaluate the performance of the models is average absolute error, E_{ave} . It makes use of the summation of the absolute errors and measures the forecast accuracy by averaging the magnitudes of the forecast errors, i.e. the absolute values of each error. This method is useful for measuring forecast errors in the same units as the original data. The E_{ave} is calculated using equation (2) :

$$E_{ave} = \frac{1}{N} \sum_{t=1}^N |Y_t - \hat{Y}_t| \quad (2)$$

where N is the total number of forecast, Y_t actual value at period t and \hat{Y}_t is the forecast result at period t .

3.2. Results

Simulation results in April 2006 for each port of each model are compared. Figure 2 shows the average absolute error of the simulation in the morning time periods (period 7 to period 11) for all stations. It is observed that, nn_Outliers outperforms the benchmark model in 8 out of 12 stations and the forecast error is closer to the benchmark model in 2 out of 12 ports. nn_Cluster obtains better forecast accuracy in 3 out of 12 ports as compared to the benchmark model while in 2 out of 12 ports, the simulation error is again very close to the error generated by nn model. It can be said nn_Outliers is a better model as compared to nn_Cluster. nn_Cluster however, may show a weaker result compared to the benchmark model, although in some ports, it performs better.

For the afternoon period (period 12 to period 17), comparing to the benchmark model, nn_Outliers shows stronger forecast capability in 7 out of 12 ports and in 3 out of 12 ports, both models perform closely well. nn_Cluster performs better in 4 out of 12 ports as compared to nn while the simulation error is very close to nn in 2 out of 12 ports. This is shown in Figure 3. Based on this outcome, nn_Outliers is considered to be a better model as compared to nn_Cluster. Figure 4 shows the comparison of the forecast error of each of the model in evening time periods (period 18 to period 22) for all stations. nn_Outliers has either superior or equally well performance as compared to the benchmark model in 9 out of 12 ports and the two models perform pretty much the same in 2 out of 12 stations. nn_Cluster obtains either better or equally well forecast accuracy in 3 out of 12 ports as compared to nn while the simulation error of nn and nn_Cluster is about the same in 2 out of 12 stations. Based on the evening demand forecasting error, both noise filtering techniques have once give the similar result in which nn_Outliers is considered to be a better model as compared to nn_Cluster. Although nn_Cluster could perform better in some ports, but the improvement is little compared to the nn_Outlier model.

Finally, the data is gathered and analyzed for the whole day period (period 7 to 22). Figure 5 shows that nn_Outliers performs either better or equally well as compared to the benchmark model, nn in 8 out of 12 ports and the two models perform closely well in 1 out of 12 ports. nn_Cluster achieves better or equally well forecast accuracy in 3 out of 12 ports as compared to nn while the simulation error is very closed to nn in 2 out of

12 ports. Based on this result, we could get a similar conclusion.

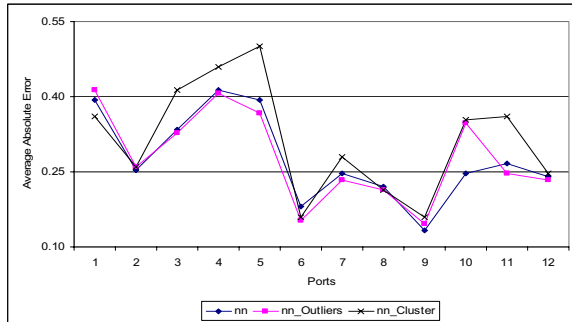


Figure 2: Models comparison of nn, nn_Outliers and nn_Cluster in morning time period

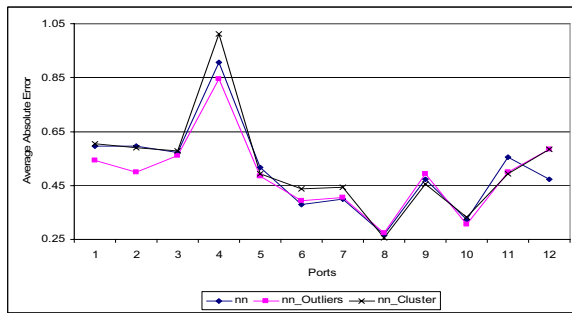


Figure 3: Models comparison of nn, nn_Outliers and nn_Cluster in afternoon time period

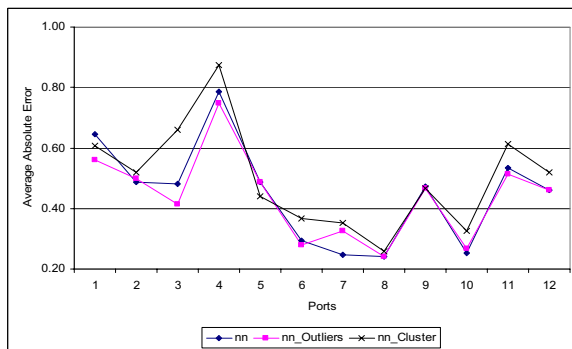


Figure 4: Models comparison of nn, nn_Outliers and nn_Cluster in evening time period

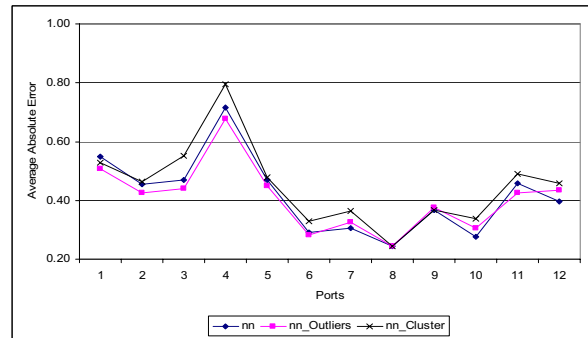


Figure 5: Models comparison of nn, nn_Outliers and nn_Cluster in whole day

3.3. Discussion

It is expected that the forecast error for nn and nn_Outliers models is close, because the same input variables are being fed into the neural network simulation model, yet, the application of outliers analysis filtering technique in nn_Outliers model has effectively removed the unusual usage in the data set. The dissimilarity of the data set is further reduced, which will enhance the ability of the model to generalize its setting on weightings and biases in order to tune to a more robust forecasting model. This explains the stronger forecast strength in nn_Outliers model. However, some of the stations may not show any improvement in the forecast and it may be because of the method being used in removal of outliers is not optimal. Since data visualization method is adopted in this analysis and it could be considered to be a subjective measure. Therefore, outlier eliminations must be handled cautiously as the incorrect removal of data will lead to a loss of important hidden information.

In nn_Cluster, instead of applying the same input variables to all time periods, it takes in different input variables selected by K-means clustering for different time periods in each station. Therefore, it is expected that the forecast performance of nn_Cluster will differ from nn and nn_Outliers. The advantage of applying clustering analysis is that inputs which are highly relevant to forecast period are identified and these specialized input variables will enhance the model simulation and the improvement is shown in certain stations that forecast error is reduced. However, sometimes, due to spontaneous human behavior, certain patterns cannot be tracked so easily and as a result, K-means clustering may not be strong enough to provide clusters of inputs that are similar to that particular time period. Hence, certain stations may not be benefited from clustering analysis filtering technique.

4. Conclusion

This is a novel study on the application of various

intelligent noise filtering techniques in the demand forecasting for the carsharing system. By applying the appropriate filtering techniques before feeding the data into the simulation model shows preliminary promising result in enhancement of the forecasting model and improvement in prediction accuracy. This finding may help improving the efficiency of the carsharing operation. From the results, it is observed that outlier analysis is an effective filtering technique to be applied while clustering analysis is a weaker filtering technique as it is strongly affected by spontaneous human behavior. One could see that nn_Cluster may have little or no improvement compared to the benchmark model. Hence, a proper choice on the filtering techniques is very crucial to ensure the effectiveness of the techniques. An appropriate choice of the filtering technique may improve the accuracy of the prediction. For future work, more testing can be done using different month of data and more filtering techniques can be tested while different forecasting simulation model can also be adopted.

5. Acknowledgement

The authors would like to thank Prof. Xu Jianxin from the Department of Electrical and Computer Engineering, National University of Singapore and Mr. Lim Jyh Shyong for their valuable comments to the study. In addition, the authors also would like to express their sincere gratitude to Honda Diracc for their substantial support throughout the study and generously fund the authors to the 14th World Congress of Intelligent Transportation Systems in Beijing, 2007.

6. References

- [1] Shaheen, S., Cohen, A.P. and Roberts, J.D. "Carsharing in North America: Market growth, current developments, and future potential", in Proceedings of the 85th Annual Meeting of Transportation Research Board, Washington, D.C., 2005.CD-ROM
- [2] Barth, M., Shaheen, S., Fukuda, T. and Fukuda, A., "Carsharing and station cars in Asia: An overview of Japan and Singapore", in Proceedings of the 85th Annual Meeting of Transportation Research Board, Washington, D.C., 2005.CD-ROM
- [3] Klatt, S, "Carsharing-an integral part of an intermodal transport system", *European Parliament-Scientific and Technological Options Assessment, briefing Note No.4*, 2001
- [4] Kek, A, *Decision support tools for carsharing systems with flexible return time and stations*, Doctor of Philosophy Dissertation, National University of Singapore, Singapore, 2006.
- [5] Britton, E. and Associates, "Carsharing 2000: Sustainable transport's missing link", *Journal of World Transport Policy and Practice*, vol. 5, no.3.
- [6] Millard-Ball, A., Murray, G., ter Schure, J., Fox, C. and Burkhardt, J., *TCRP Report 108, Car-sharing: Where and how it succeeds*, Transportation Research Board, Washington D.C. 2005
- [7] Barth, M. and Todd, M. "User behaviour evaluation of an intelligent shared electric vehicle system", in Proceedings of the 80th Annual Meeting of Transportation Research Board, Washington, D.C., 2001.CD-ROM.
- [8] Barth, M. and Todd, M. "Performance evaluation of a multi-station shared vehicle system", in Proceedings of the 4th IEEE International Conference on Intelligent Transportation Systems, IEEE, pp. 1218-1223.2001.
- [9] Allouche, J.F., benejam, E., Massot, M.H. and Parent, M., "Praxitele: Preliminary results from the Saint-Quentin experiment", in Proceedings of the 78th Annual Meeting of Transportation Research Board, Washington, D.C., 1999.CD-ROM
- [10] Honda Diracc, <http://www.hondadiracc.com.sg>. Accessed February, 2006.
- [11] Kek, A., Cheu,R.L. and Xu,J.X. "Trip forecasting models for intelligent community vehicle systems", in Proceedings of the 84th Annual Meeting of the Transportation Research Board, Washington, D.C., 2005. CD-ROM.
- [12] Cheu, R.L., Xu, J., Kek, A., Lim, W.P. and Chen, W.L., "Forecasting of shared-use vehicle trips using neural networks and support vector machines", in Proceedings of the 85th Annual Meeting of Transportation Research Board, Washington, D.C., 2005.CD-ROM
- [13] Kantardzic, M, *Data mining: Concepts, models, methods, and algorithms*, John Wiley and Sons Inc., USA, 2003.



Hooi Ling KHOO She is currently a PhD candidate in NUS. She obtained her Bachelor degree (Civil Engineering) and Master Science Engineering from University of Malaya, Malaysia. She has published a few papers in the international journals and conferences.



Chau Ha FUNG graduated from University of Wisconsin, Madison with a Master of Science, Electrical Engineering degree. Her research area is in data mining and model simulation. She is currently working in equity research area.



Dr. Qiang MENG obtained his PhD degree from the Hong Kong University of Science and Technology in year 2000. His expertise is in transportation and logistics system optimization and modeling. He has published over 20 papers in international journals and conferences.



Dr. Der-Horng LEE graduated from the University of Illinois in year 1996 and won the TR100 Honouree in year 2002. His research interest is in microscopic simulation, intelligent transportation systems and transportation modeling and analysis.

Received date: December 10 2007

Received in revised form: March 15 2008

Accepted date: March 16 2008

Editor: Keping Li