Link-based Emission Model for Eco Routing

Jianchu Kang[1], Tao Ma[1], Fajin Ma[1], Jian Huang[2]
State Key Laboratory of Software Development Environment
Beihang University
Beijing, China
[1]{kang, tao_m, mafajin}@nlsde.buaa.edu.cn
[2]{huangjian}@buaa.edu.cn

Abstract—**the tremendous growth of vehicles has led to severe problems for energy and the environment. More and more research work on eco driving has been made to devote the fuel efficient intelligent driving. This paper works on the eco routing, which is one of important aspects of eco driving, and which is to find some more eco-friendly routes for the driver in terms of minimum fuel consumption and vehicle emissions. The paper first abstracts six primary driving patterns from large amounts of real data which represent, respectively, the velocity, acceleration and engine stress of the vehicles driving on some links. And then, by associating the six driving patterns with some other properties of the links, like traffic congestion, etc., a link-based driving pattern classifier is established, which is applied to calculate and estimate the fuel emission characteristics for each relevant link. At the end of this paper, the effectiveness of the distance priority routes, time priority routes and eco routes are compared, the eco route one demonstrates its superiority.**

Keywords- vehicle emission, driving pattern, link-based emission model, eco routing

I. INTRODUCTION

As the continuously increasing of vehicle miles of travel, the transportation activities are consuming more and more energy resources which lead to the aggravation of air pollution. It has been estimated that transport accounts for about 21% CO_2 emission globally although there are considerable differences between global regions [1]. According to China Vehicle Emission Annual Report 2010, over fifty million tons of pollutions are emitted directly from ground transport in one year. Vehicles were the main contributors which subscribed more than 70% CO, HC, and more than 90% NOx, PM [2]. Reducing vehicle emissions has already been one of the most significant issues today.

This problem is to be addressed, not only by improving vehicle efficiency and developing alternative fuels, but also by reducing vehicle miles traveled and making roadway travel more efficiently [3]. Recently a new navigation method called "eco routing" is developed [1, 3]. Its purpose is to find a route with the least amount of vehicle emissions. However, the vehicle emissions are affected by many factors such as the types of autos, the road condition and driving pattern of the drivers, etc. Therefore, how to estimate the quantities of the vehicle emissions under various conditions becomes a crucial issue of eco routing research.

In this paper, a new link-based emission model is constructed based on the analysis of driving behavior and the vehicle-based emission model. This new model will give strong support to eco route systems. In the following, section II gives the technique background of link-based emission model; section III presents the modeling for link-based emission; in section IV, some experiment results are showed, which demonstrates the validity of the link-based emission model; and section V provides the conclusions and future work.

II. TECHNIQUE BACKGROUND

In order to design the link-based emission model properly, the following three fundamental techniques are involved, that are the real-time traffic measurement, driving activity analysis and vehicle emission model.

A. Real-time Traffic Measurement

Though the network of roads is static, the traffic should be viewed dynamically, which is constituted by the flows of vehicles, and which varies quite a lot due to some subjective or objective reasons. The traffic state acquisition is viewed as part of the fundamental research work of Intelligent Transportation System (ITS). Experiencing profound growth, some basic technologies, such as real-time traffic measurement, are becoming more and more mature. There have already been several successful traffic measurement systems such as California's Freeway Performance Measurement System (PeMS) in the U.S., Vehicle Information and Communication System (VICS) in Japan, etc.

Beijing's Dynamic Traffic Information Service System is such a kind of traffic measurement system. It utilizes more than 140,000 float cars as mobile sensors together with loop detectors to capture the traffic conditions and publishes the traffic information in every 5 minutes to the drivers. The published information includes: traffic congestion level, travel time, state and emergency of each link. This information is utilized in the research to capture the actual road conditions which plays an important role in vehicle emission.

978-1-61284-671-2/11/$26.00 ©2011 IEEE

B. Driving Activity Analysis

Besides the traffic condition, the driving pattern, such as the driving speed and accelerate, is also a key factor for vehicle emissions. In the view point of fuel saving and increasing traffic safety, many researches have proposed their ways to rank or classify the driving behaviors of the drivers according to the factors of speed, acceleration and engine stress, etc. [4, 5, 6]. The impact of different driving patterns on vehicle emissions is also investigated. It has been found that the driving style and fuel consumption behave differently under various conditions, e.g., in moderate driving accelerations dominate the fuel consumption, while, in high speed driving the velocities have greater impact [7].

Therefore, an accurate estimate of the driving patterns at different situations is necessary for the link-based emissions calculations. And the same time, several driving parameters should be selected to represent the driving pattern. In this paper, a large database containing more than two hundred hours of real traveling data collected in recent two years is built, which is supposed to be used in analyzing the driving behaviors of the drivers.

C. Vehicle Emission Model

Vehicle emission model may be viewed as a vehicle emission simulator, which can be used for estimating the emissions for mobile sources (such as cars, trucks and buses) covering a broad range of pollutants.

To analyze the environmental impacts from the ground transport, the U.S. and European have already developed several emission models based only on their own emission data. In many cases, these models can lead to some errors in emissions estimates. To solve this problem, a new one called the International Vehicle Emissions (IVE) Model [8] has been built which is developed jointly by researchers at the International Sustainable Systems Research Center and the University of California at Riverside, and was applied successively in Mexico City, Santiago, Beijing, and 19 other locations.

The basis of the emission prediction process of the IVE model is to apply a base emission rate with a series of adjustment factors to estimate the amount of pollution. These factors mainly refer to temperature, humidity, fuel quality and driving behavior. All but the driving related factors can be directly derived from relevant statistics. Since driving behaviors are affected not only by the driving style but also by the road condition, vehicle type and driving environment, it is difficult to capture the actual driving patterns of each vehicle. In IVE model, the various driving patterns are merely represented by a vehicle specific power distribution calculated from several hours of driving data, which is too general for the estimates of link-based emissions. Hence the more precise driving patterns are needed to accurately achieve the link-based emissions. The power-based driving factors are specially studied and applied by this paper mentioned in section III.

III. LINK-BASED ENERGY/EMISSION MODELING

This section includes the following three parts of work. The first one is to obtain the driving patterns by analyzing and clustering with the large amounts of real driving data; the second one is to construct the link-based classifier that will be used to determine which kind of driving pattern a vehicle on a specific link behaves in; the third part of work is to calculate the link-based emissions by applying together the IVE model with the corresponding driving pattern.

A. Obtaining Typical Driving Patterns

As mentioned above, this paper has collected over two hundred hours of driving data to support the analysis of the driving activities. In generally, the data related to the driving behaviors are first processed by means of noise cleaning and map matching. Then, the noise free dada of the vehicles are split into groups, corresponding to the links they matched on the map; and the power distributions of the vehicles are calculated for each group. Finally, the X means algorithm [12] is used to achieve the typical driving patterns.

1) *Data* pretreatment

The driving data is pretreatment mainly through the noise cleaning and map matching.

Because of the positioning failure of GPS, the driving data of devices contain some abnormal values. Though the error records are few, the noises will still affect the following calculations significantly due to their enormous values. Hence all of the data beyond the normal speed range or the geography scope are filtered out from the database.

Besides, owing to the positioning offset of GPS, many driving records about the vehicles do not fall on the road, but the record trajectory's direction remains almost correct, as shown in figure 1. Map matching means to find the nearest link on which the particular vehicle is driving, and, at the same time, the premise should be obeyed, which ensure the vehicle is driving on the coherent direction. Figure 2 illustrates the procedure. To begin with, the map matching program searches all the links within 50 meters of the current records. Then it is checked whether the link and track directions of records are inconsistent, i.e. the deviation of both directions are too large. All of the inconsistent links are removed. Finally the link records with the minimum offset are chosen to be the matched link.

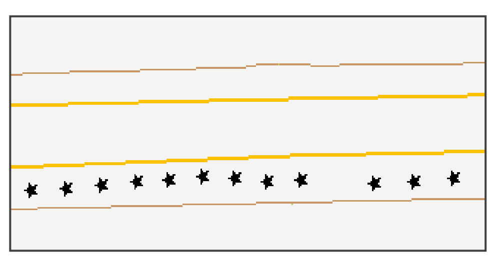

Figure 1. Offset of the GPS records

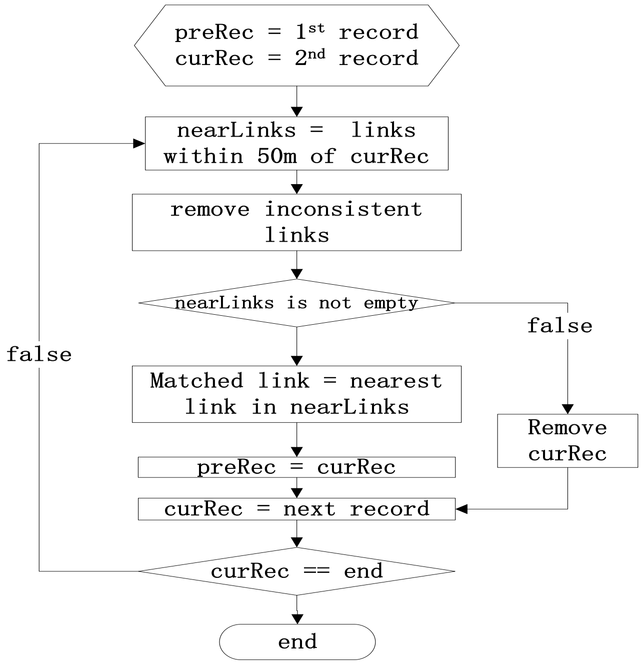

Figure 2. Map matching procedure

2) Characterizing Driving Pattern

As mentioned in section II, the vehicle emissions can be achieved by applying a base emission rate with fuel quality, driving behavior, geographic and climatic factors. And the key issue is to precisely acquire the driving patterns of the vehicles under different driving conditions. However, driving pattern includes many parameters such as velocity, acceleration, engine rpm, etc. and each of them affects the vehicle emissions differently. For this reason, the most important driving parameters should be chosen at first, so as to characterize emissions as a function of these parameters.

According to the conclusions of IVE model, it was determined that the single most important parameter for determining emissions is the vehicle specific power, which is derived from the instantaneous velocity and acceleration [9]. The equation for VSP (KW/ton), first developed for this application by Jimenez Palacios, is shown in (1) [10].

$$VSP = v[1.1a + 9.81(\mathrm{atan}(\sin(grade))) + 0.132] + 0.000302v^3 \quad (1)$$

where v = velocity (m/s)
a = acceleration (m/s2)
grade = (ht=0 – ht=-1)/v(t=-1to0)
h = Altitude (m)

The VSP based emissions estimates perform quite well for CO_2, but improvements in predictive power for other emissions such as CO, HC, NO_x, and NH_3 may be achieved through the addition of one or more dimensions [11]. Another parameter called engine stress is used in addition to VSP. Engine stress is shown to correlate best to vehicle power load requirements over the past 20 seconds of operation and implied engine RPM (Equation. 2, Table I) [10].

$$\text{Engine Stress} = \text{RPMIndex} + (0.08\ \text{ton/KW}) * \text{PreaveragePower} \quad (2)$$

where $\text{PreaveragePower} = \text{Average}(VSP_{t=-5sec\ to\ -25\ sec})$ (KW/ton)
$\text{RPMIndex} = \text{Velocity}_{t=0}/\text{SpeedDivider}$
Minimum RPMIndex = 0.9

TABLE I. CUTPOINTS USED IN RPMINDEX CALCULATIONS

Speed Cutpoints (m/s)		Power Cutpoints (KW/ton)		Speed Divider (s/m)
Min	*Max*	*Min*	*Max*	
0.0	5.4	-20	400	3
5.4	8.5	-20	16	5
5.4	8.5	16	400	3
8.5	12.5	-20	16	7
8.5	12.5	16	400	5
12.5	50	-20	16	13
12.5	50	16	400	5

Consequently, VSP and engine stress are chosen to characterize the driving patterns. To consist with IVE model [10], the VSP and engine stress are divided into 20, 3 intervals, respectively. So a total of 60 VSP/stress categories are used, and each driving pattern is represented by a distribution of VSP/stress categories.

3) Distribution Records Cluster

All of the driving data are split into groups according the link they matched to. For each group the VSP/stress distributions are calculated. Finally about 20000 distribution records are generated from the driving data.

As the link-based distribution record has 60 numerical attributes and all of the attribute values are normalized, it just suit to X-means cluster algorithm [12]. The X-means algorithm is used to find k clusters so as to minimize the with-in cluster sum of squares. The maximum number of iterations is set to 1000. The range of clusters' number is 3 to 8, i.e. the maximum acceptable number of cluster is 8 and the minimum number is 3.

Totally, six clusters are generated by the X-means algorithm, as shown in figure 3. Generally speaking, the VSP distributions can be classified into two categories. The first one includes the previous three clusters which represent the situations of heavy traffic where most of time of the vehicles is wasted on waiting or crawling. However, there are some subtle differences among these three ones. The average speed of the first cluster is the lowest, and over 80% of time is spent on waiting (the 12th interval); The speed of second one is a little faster, and the high engine stress part (40th -59th interval) almost has none of data; The third cluster is the fastest one in this category, and has a few high engine stress data. The second category includes the rest three clusters which represent the driving patterns in better traffic conditions. Though the low speed section still contains about 40% data, the other sections have obviously more data than the first category. The 6th cluster even has about 5% high engine stress data which indicates the traffic condition is good enough to enable the high speed driving.

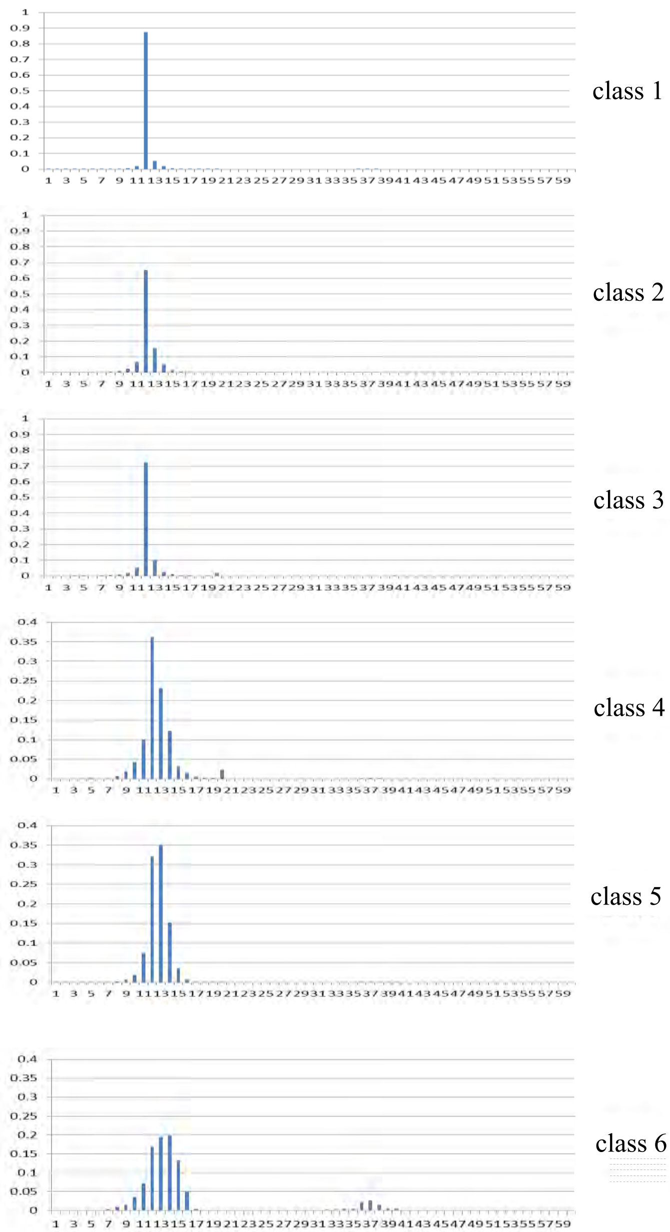

Figure 3. Six cluster centers of VSP distributions

B. *link-based Classifier*

This part explicitly explains the approach to construct the link-based classifier. First the link's attributes, which is utilized by classifier algorithm, are introduced. Then the link-based classifier is developed using PART classifier algorithm.

1) Link attributes

To some extent, the more attributes a training dataset has, the more accurate the classifier will be. But many attributes of the links are too difficult to acquire. Currently there are only 5 attributes available and used in this research, they are, the kinds, length, width, fee and average speed of the roads involved, which will be described in details as following.

The kind of link denotes road's level and properties. According to their capacities, speed limits and functionalities, links are classified into seven levels, express way, urban freeway, national highway, provincial highway, prefectural highway, country way, and others (secondary road). For each level, links are classified into five categories, road with separate up and down, ramp, roundabout, parking area and normal road.

The link's length is another important factor which indicates the density of junctions. The precise length of each link is achieved through the geographic database.

The width of link indirectly reflects the level and quality of the road. In this research the width is discretized into four values, 15, 30, 55 and 130, represent the width below 3 meters, 3-5.5 meters, 5.5-13 meters and over 13 meters respectively.

The links have four kinds of toll states: unknown, toll road, free, free part in toll road. Different toll states indicate different states of maintenance which could also impact driving pattern to some extent.

The average speed of links reflects the current traffic conditions which is a dominant factor of driving pattern. For instance, congestion usually causes the frequent stop-and-go driving; while in better ones, vehicles will be driven more smoothly.

2) Link-based classifier

To acquire the relations between links and typical driving patterns, classification techniques are utilized in this research. First of all, a training dataset is generated from the link-based VSP/stress distribution records alone with attributes values of the corresponding link. Using this dataset, several topping classification algorithms are tested on their accuracy, complexity and interpretability. By comparing the results, PART is selected for its better accuracy and interpretability. This algorithm builds a partial C4.5 decision tree in its each iteration and makes the best leaf into a rule to constitute the final decision list [13]. And C4.5, developed by Quinlan in 1933 [14], is well known for its simplicity and accuracy. It is a decision tree based algorithm, using normalized information gain as criterion.

The confidence factor used for pruning is set to 0.25 and the minimum number of instances per leaf is set to 2. The final decision list totally has 137 rules and table II only displays part of them due to the limited space. The key-value pairs before colon are antecedents of the rule and the value behind the colon is the consequence, i.e. the index of driving pattern corresponding with this rule. The fraction within the parentheses implies the coverage and accuracy of the rule.

Numerator and denominator denote the correct and wrong classification number.

TABLE II. PARTIAL DECISION LIST

average speed <= 1.538952: 1 (563.0/2.0)
average speed <= 2.260017 & length > 0.202: 1 (77.0/3.0)
average speed > 21.256657: 6 (105.0/3.0)
average speed > 16.644048 & kind = 102 & width = 130: 6 (65.0/24.0)
average speed > 11.234596 & kind = 102 & width = 130: 5 (75.0/19.0)
average speed > 18.005991 & length <= 0.922: 6 (60.0/13.0)
average speed <= 4.467196 & kind = 601 & toll = 0: 1 (9.0/2.0)
average speed > 11.234596 & kind = 202: 5 (10.0/2.0)
average speed > 11.336623 & toll = 3: 5 (296.0/99.0)
average speed <= 4.467196 & kind = 5: 3 (5.0)
…

C. Calculation of link-based emission

As introduced in section II, the quantity of emission is calculated by applying base emission ratio with a series of adjustment factors. The link-based emission is calculated though two steps.

In the first step, two types of relatively static factors are calculated. One is local factor the other is vehicle related factor. The local factor contains four kinds of variables, the ambient temperature, the humidity, the altitude and the inspection/maintenance categories. The vehicle related factors include the fuel quality variables and the vehicle technology variables. Most of the adjustment factors are acquired from the result of IVE's previous validation program held in Beijing in 2005 [15].

In the second step, each link is classified as one of the six driving patterns by the link-based classifier. And the driving adjustment factor of each pattern is calculated using equation 3.

$$\text{DriAdjustmentFactor} = \sum_{i=1}^{60} C_i * P_i \qquad (3)$$

where C_i is the adjustment factor for the i^{th} category.

P_i is the adjustment of the i^{th} category

The final link-based emission is achieved by multiplying base emission ration by these three kinds of adjustment factors.

IV. EXPERIMENT RESULT

A. Evaluation of link-based classifier

In the link-based emission model, the link-based classifier is viewed as the essential component. Its accuracy will directly affect the model's performance. To prove its accuracy and robustness, the classifier is validated on another driving dataset which is generated from 30 hours of driving records. All of the records are processed by the same way as explained in the section III. And totally, 2951 instances are generated. All of them are classified by the link-based classifier. As a result, 2446 instances are correctly classified and the rest 505 are incorrectly classified. Table III shows the confusion matrix. It can be found that instances with class 1 or class 2 are more likely to be falsely classified. The main reason for this is that these two classes have too few instances; so the classifier does not have sufficient data to provide accurate rules for these two classes. The deficiency of link's attributes is another reason for incorrect classification. Currently only five link's attributes are available. The more detail information about the roads is not stored in the existing database. Maybe some additional attributes will further enhance the classifier's accuracy.

TABLE III. CONFUSION MATRIX

classified as	**1**	**2**	**3**	**4**	**5**	**6**
1	7	0	1	12	0	0
2	0	2	0	1	17	11
3	0	0	249	151	7	0
4	2	0	49	626	95	0
5	0	2	3	46	689	59
6	0	0	0	1	48	873

B. Emissions comparison for different route choices

The travel distance and time are two of the most common criteria for car navigation. Emission reduction is usually achieved as a byproduct. But, in many cases, the distance priority route or the time priority route is not the optimum solution considering of vehicle emissions. To demonstrate the existence and evaluate the emission reduction effect of the ecological route, a comparative experiment has been conducted in which the effectiveness of the distance priority route, the time priority route and the eco-route are compared as following.

As Figure 4 shows, the tested trip is from Beihang University to ShouDi shopping center, denoted by the blue and red point respectively. And the blue, purple, green lines represent the time priority route, distance priority route and eco-route. The travel time, length and CO2 emission corresponding to the three types of routes are illustrated in Figure 5. It can be found several instructive phenomena. First, the time minimization route chooses more smoothing links to save the travel time. Though the route is much longer, it avoids the congestions. Therefore the travel time is reduced. But it causes much more emissions at the same time. Second, the shortest route selects the most direct links to the destination no matter what the driving condition is. Consequently neither the travel time nor the emission is the optimum. Third, the eco-route is a bit similar to the shortest route. To reduce emissions, the eco-route makes a tradeoff between travel time and distance. The increment in either time or distance will lead to more emissions.

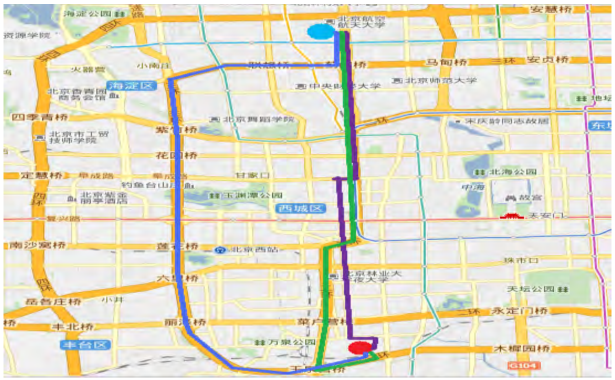

Figure 4. Three kinds of routes: blue-time priority route, green-eco route, purple-shortest route

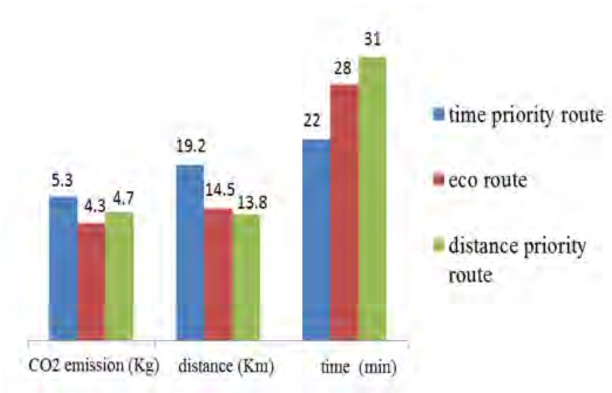

Figure 5. Traveling time, distance and CO2 emission of each route

V. CONCLUTIONS

This paper presents a link-based emission model that can be used for each link to estimate the quantities of emissions produced by vehicle passing through it. First of all, six typical driving patterns have been abstracted from a large amounts of real data. Afterwards, a link-based driving pattern classifier is developed to accurately estimate the driving patterns on various kinds of roads. Finally, the link-based emissions are calculated according to their corresponding patterns. As shown in section IV, the link-based emission model may be conveniently applied for those navigation systems to provide the eco routes for drivers. And the experiments have demonstrated the link-based classifier's accuracy and reliability, as well as the link-based model's effectiveness.

Although the link-based emission model is fairly effective, some more works are needed to enhance its performance. On the one hand, the information about the road network is too limited. More attributes are required to sufficiently capture the relation between roads and driving patterns. On another hand, the eco routing system could be more intellegent by containing a traffic predictive model, so as for it to be able to provide optimal routes considering not only the current but also the future traffic conditions.

ACKNOWLEDGMENT

This research was supported by Megalopolis Advanced Transportation Operate Coordinate Command Platform Program (No.2010ZX01045-001-009-1) and the fund of the State Key Laboratory of Software Development Environment (No. SKLSDE-2009ZX-09)

REFERENCES

[1] Eva Ericsson, Hanna Larsson, Karin Brundell-Freij, "Optimizing route choice for lowest fuel consumption – Potential effects of a new driver support tool", Transportation Research Part C, Volume 14, Issue 6, December 2006, pp 369-383

[2] ministration of environmental protection the people's republic of China, "China Vehicle Emission Annual Report", 2011, http://www.vecc-mep.org.cn/index/20101110nianbao.pdf

[3] Kanok Boriboonsomsin, Matthew Barth, Weihua Zhu, Alexander Vu, "ECO-Routing Navigation System based on Multi-Source Historical and Real-Time Traffic Information", IEEE ITSC2010 Workshop, 2010

[4] Z. Constantinescu, C. Marinoiu, M. Vladoiu, "Driving Style Analysis Using Data Mining Techniques", Int. J. of Computers, Communications & Control, Vol. V (2010), No. 5, pp. 654-663

[5] Rahul Mundke, Sachitanand Malewar, Kavi Arya, "Use of Data Recorder for Driver Rating", SAE International, 2006

[6] M.Tech Dissertation, "DRIVER RATING SYSTEM USING DATA RECORDER", Kanwal Rekhi School of Information Technology , Master's degree theses, 2005

[7] Irene Michelle Berry, "The Effects of Driving Style and Vehicle Performance on the Real-World Fuel Consumption of U.S. Light-Duty Vehicles", Massachusetts Institute of Technology, Master's degree theses, 2010

[8] Nicole Davis, James Lents, Mauricio Osses, Nick Nikkila, Matthew Barth, "Development and Application of an International Vehicle Emissions Model", Transportation Research Board 81st Annual Meeting, January 2005, Washington, D.C.

[9] International Sustainable Systems Research Center, Attachment C Characterizing Emission Variations due to Driving Behavior from On-Road Vehicles, www.issrc.org/ive/

[10] Jiménez-Palacios, "Understanding and Quantifying Moto Vehicle Emissions and Vehicle Specific Power and TILDAS Remote Sensing", Massachusetts Institute of Technology Ph.D theses, 1999

[11] International Sustainable Systems Research Center, Attachment B Development of the Correction Factors for Use in the IVE Model, www.issrc.org/ive/

[12] Dan Pelleg, Andrew W. Moore, "X-means: Extending K-means with Efficient Estimation of the Number of Clusters", Seventeenth International Conference on Machine Learning, pp 727-734, 2000.

[13] Eibe Frank, Ian H. Witten, "Generating Accurate Rule Sets Without Global Optimization", Fifteenth International Conference on Machine Learning, pp 144-151, 1998.

[14] Quinlan, J. R, "C4.5: Programs for Machine Learning", Morgan Kaufmann Publishers, 1993

[15] Huan Liu, He Chunyu, James Lents, Nicole Davis, Mauricio Osses, Nick Nikkila, "Beijing Vehicle Activity Study", IVE report, www.issrc.org/ive/, 2005