An Integrated Map Matching Algorithm for GPS-Based Freeway Network Traffic Monitoring

Wei Wang *
Ph. D. Candidate
School of Transportation, Southeast University
Nanjing, Jiangsu, China 210096
Department of Civil and Environmental Engineering,
University of Wisconsin-Madison,
Madison, WI 53706
Phone: 86-25-83795528
E-mail: wangseu@gmail.com

Jing Jin
Ph. D. Candidate
Department of Civil and Environmental Engineering
University of Wisconsin-Madison
Madison, WI 53706
Phone: 1-608-262-2524
E-mail: jjin2@wisc.edu

Bin Ran
Professor
Department of Civil and Environmental Engineering
University of Wisconsin-Madison
Madison, WI 53706
Phone: 1-608-262-0052
E-mail: bran@wisc.edu

Xiucheng Guo
Professor
School of Transportation, Southeast University
Nanjing, Jiangsu, China 210096
Phone: 86-25-83795427
E-mail: seuguo@163.com

* Corresponding Author

5600 words + 7 figures (250 × 7) = 7350 words

Submission date: July 31, 2009

ABSTRACT

This paper proposed an integrated map matching algorithm for freeway network traffic monitoring system using GPS probe data. The algorithm is novel in three aspects. First, it is designed to map massive GPS positioning points in a large-scale traffic network for the purpose of dynamic traffic state monitoring, while most the existing map matching algorithms are designed for in-vehicle navigation. Second, the modified N-shortest path algorithm is embedded to find reasonable routes between GPS positioning points efficiently. Third, the algorithm uses fuzzy logic inference system with innovative fuzzy rules to determine the actual traveling route and physical location for each probe vehicle. The algorithm is evaluated using field GPS data sets collected in Los Angeles. This paper uses a new evaluation strategy that has not been used in previous literatures. In addition to checking several map matching cases, the speed detection results of GPS probe implementing the proposed map matching algorithm are compared with the speed data from the loop detecting system. In this way, the map matching results are converted to a dataset that can be evaluated using ground truth data. The results demonstrate the effectiveness and robustness of the proposed algorithm for mapping mass GPS positioning data onto complex freeway network, as well as producing accurate speed detection results.

KEYWORDS: Global Positioning System (GPS); Map Matching; Freeway Traffic Monitoring; Shortest Path; Fuzzy Logic

INTRODUCTION

Traffic congestion has been a national issue. According to The Texas Transportation Institute (2000), the cost of traffic congestion in the 75 largest U.S. metropolitan areas is 67.5 billion. Advanced Traveler Information System (ATIS), a key component of Intelligent Transportation System (ITS), is one of the mitigating solutions to the congestion problem. ATIS can improve the road network efficiency and safety by providing the traveler and traffic control center with real-time traffic information. Nowadays the loop detectors and video cameras still are the major data sources of ATIS, but as the market penetration of Global Positioning System (GPS) equipments in vehicles increases over the year GPS probe data has become an important emerging data source for ATIS.

The key procedure of GPS probe-based traffic monitoring system is the map matching algorithm. The purpose of map matching*(1)* is to find the actual road segment on which each probe vehicle is driving and project each GPS positioning point to physical position on a link within reasonable processing time. The traditional map matching algorithms are designed for in-vehicle navigation, while in this study the proposed map matching algorithm is used to detect traffic states throughout a road network. The different design purpose of map matching can cause different characteristics (e.g. input data) and different requirements (e.g. accuracy, computation cost etc.) for the algorithms. For traffic monitoring, the following things must be taken into account carefully for the proposed map matching algorithm:

- Traffic monitoring needs the traffic information for a specific area to be aggregated in a given time interval, such as the average speed for each road segment every five or fifteen minutes. This characteristic allows longer processing time than traditional map matching methods. We do not need to process each new data point immediately when it comes into database; instead, we can execute the map matching procedure periodically and process all the GPS data points of the current time interval at one time.
- To provide accurate real-time traffic information, the proposed map matching algorithm must have the ability to handle complex network geometries, such as frontage roads, different types of interchange, and so on.
- The map matching algorithm proposed here is for entire freeway network traffic state monitoring. The data set needs to be processed is massive. Therefore efficiency of the algorithm is one of the objectives.

LITERATURE REVIEW

As stated above, existing map matching algorithms can be classified into two categories, the first one is for in-vehicle navigation *(2-22)*, the other one is for network traffic state monitoring*(23, 24)*.

Map Matching Algorithm for In-vehicle Navigation

Bernstein et al.(1996)*(2)* classified in-vehicle map matching algorithms into two categories: 1) geometric analysis based algorithms, and 2) geometric and topologic analysis based algorithms. Quddus et al.(2007)*(1)* added another two categories: probabilistic analysis based algorithms and advanced algorithms.

Geometric analysis based algorithms

A geometric analysis based map matching algorithm makes use of only geometric information provided by digital road network *(2, 4)*. This method does not consider whether the candidate links are connected to each other or not. The commonly used geometric analysis based algorithm is called the point-to-point method *(2, 4)*, which matches each GPS point to the closest "node" or "shape point" of the network. This method is easy to implement and the computational cost is low, but its accuracy is hard to control because the algorithm depends heavily on the number of shape points of each link for link determination. .

Another geometric analysis based algorithm is the point-to-curve approach *(2, 4, 19, 25)*, this approach matches the GPS positioning point to the curve with shortest perpendicular distance. This method is better than point-to-point method. However, it does have several disadvantages that make it inappropriate in practice. For example, the result is unstable at high density region of a roadway network, where a slight change of GPS location can cause completely different or contradictory results.

The other way is to calculate the distance between the trajectory and the piecewise linear curve, the curve with smallest distance is selected as the segment which the vehicle is travelling on. This method is known as curve-to-curve method *(5, 12)*. This method is quite sensitive to outliers and depends on preliminary results from point-to-point matching, and therefore sometimes giving unexpected results*(1)*.

Geometric and topologic analysis based algorithms

Topology information, referring to the connectivity of the network elements, can reduce the set of "candidate" points or curves dramatically. These map matching methods using both geometric and topology information is known as geometric and topologic analysis based method *(2, 7, 10, 14, 17, 19, 22, 26)*. Bernstain et al. (1996)*(2)* is the first one who proposed to use topology information in map matching, and presented methods to improve point-to-curve and curve-to-curve matching approaches. Greenfeld (2002)*(7)* and Quddus et al. (2003)*(10)* made use of topology information for their research independently, and developed different weighting schemes to enhance the map matching result.

Probabilistic analysis based algorithms

The probabilistic analysis based algorithm defines a confidence region around a GPS sample point. The confidence region is then superimposed on the road network to identify the road segment on which the vehicle is travelling. If a confidence region contains a number of segments, then the evaluation of candidate segments are carried out using heading, connectivity, and closeness criteria*(1)*. Ochieng et al. (2004)*(14)* developed a probabilistic analysis based map matching algorithm for vehicles travel through junction areas. Blazquez et al. (2005)*(17)*, Quddus et al. (2006)*(19)* used this algorithm to create "buffer" surround each GPS point.

Advanced algorithms

Advanced map matching algorithms are those that use more refined concepts from different research fields, for example Kalman filter or Extended Kalman filter*(3, 20)*, Particle filter*(8)*, Interacting multiple model*(9)*, Dempster-Shafer (D-S) theory of evidence*(11)*, Fuzzy logical model*(19, 27, 28)*, Believe function*(21)*, Multiple hypothesis technique*(6)*.

Map matching algorithm for network traffic state monitoring

Among the published map matching related papers, only two of them focused on the algorithms for road network traffic state monitoring, which are Marchal et al. (2005) *(23)* and Pan et al. (2007)*(24)*. In the algorithm proposed by Marchal et al. (2005), only GPS coordinates and the network topology were considered as input. They use the distance between a point and a road segment, score of a path (base on how many links of the path are projected by the GPS points) and path connectivity as indexes to decide on the correct link. If the GPS heading information and travel cost for each candidate path are considered in this algorithm, more reliable results may be generated. Pan et al. (2007) introduced a method to collect travel time and delay for arterials using GPS data. They developed a geometric and topology analysis based map matching algorithm for post-trip GPS data processing. To determine the actual position of the vehicle, this algorithm developed three criteria to compose a weighting scheme: distance between a point and a roadway link, similarity of roadway direction and vehicle heading, and network topological characteristics.

METHODOLOGY

Definition of Data Inputs – Digital Road Network Data and GPS Data

The digital road network used in this research follows the planar model. In the planar model, a network representation (N) of a finite road system (or set), consists of a set of curves in R^2. Each of curve is called an "arc" (or "link") representing one road segment *(2, 4, 7, 10)*. Assuming each arc is piecewise linear, any arc $A \in N$ is composed of a finite sequence of points (A^0, A^1, A^2, …, A^{na}). Each of them is in R^2. All points are classified into two categories, nodes and shape points. A node (i.e. A^0 and A^{na}) is a point at which it is possible to move from one arc to another or a point at which an arc terminates/begins*(4)*. A shape point refers to one of those inner points in an arc (i.e. A^1, A^2, …, $A^{n(a-1)}$). Each arc in N and its element points (nodes and shape points) have some identifying attributes. For nodes and shape points, the attributes are latitude/longitude coordinates that identify the spatial location. The topology and geometry attributes of each arc (road segment) are determined by the nodes within the arc (i.e. start/end node, shape points). The additional attributes (i.e. road name, arc length, speed limit etc.) of each arc needs extra inputs and are independent from attributes of nodes.

GPS is a satellite-based radio-navigation system owned and operated jointly by the US department of Defense (DOD) and department of Transportation (DOT)*(14)*. At a finite number of points in time, denoted by {0, 1, 2, …, T}, GPS devices can provide a sequence of probe vehicle's estimated positions, named as positioning points, {P^0, P^1, P^2, …, P^T}. Generally, the GPS positioning data has the following basic attributes: latitude/ longitude coordinates, heading, timestamp and estimated speed. Some GPS positioning data has more attributes like altitude, number of satellite in view etc. The latitude/longitude coordinates describe the current physical location of the probe vehicle with a predicted horizontal accuracy of 13m (95%) (US DOD, 2001)*(19, 29)*. The heading is the angle between the direction that the probe vehicle is pointing to and true north. Timestamp indicates the time when GPS device reports the positioning point P^t. The estimated speed is calculated by Doppler method and represents the spot velocity of the probe vehicle. To make the proposed map matching algorithm more compatible, this research only considers the basic attributes.

The Proposed Map Matching Algorithm

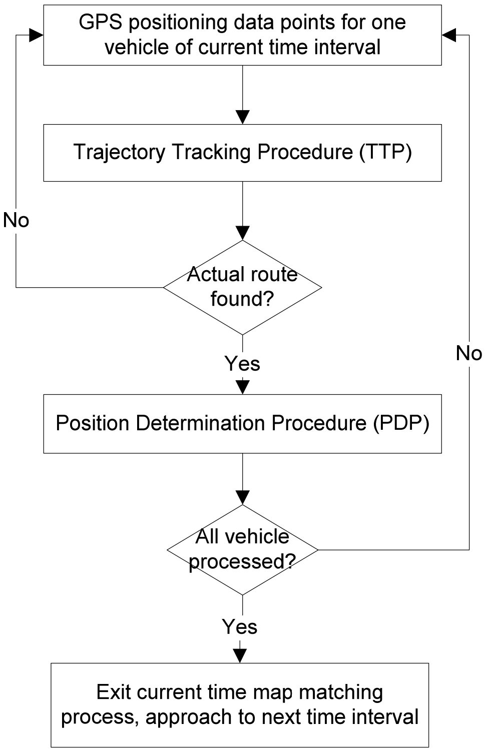


FIGURE 1 Framework of the proposed map matching algorithm.

The map matching algorithm proposed in this study is for dynamic traffic state monitoring in a complex freeway network. The analyzing period is divided into equal time slots by given time interval (usually five or fifteen minutes, depending on the purpose of the traffic sate monitoring). In each time slot T, the algorithm tracks the trajectory of each probe vehicle, selects the actual route on which each probe vehicle is driving, and finally finds the link and physical location each GPS positioning point should be matched to. The proposed algorithm does not consider the historical vehicle trajectory in previous time slot (T-1), this will make reduce the difficulty and computational cost when implementing this algorithm.

As shown in FIGURE 1, the proposed map matching algorithm has two major stages: Trajectory Tracking Procedure (TTP) and Position Determination Procedure (PDP). The purpose of TTP is to find the exact route on which each probe vehicle is travelling, and PDP is to locate the GPS positioning point on its actual position on the link.

Trajectory Tracking Procedure (TTP)

Details of Trajectory Tracking procedure (TTP) is shown in FIGURE 2. For current time slot T, we select all the positioning points of the target probe vehicle i, noting them as $\{P^0_{T,i}$, $P^1_{T,i}$, $P^2_{T,i}$, …, $P^t_{T,i}\}$. If there is only one positioning point for vehicle i, the current vehicle will be skipped and the algorithm will continue to process the next vehicle (i+1) because the algorithm needs at least two positioning points for the subsequent route choice analysis.

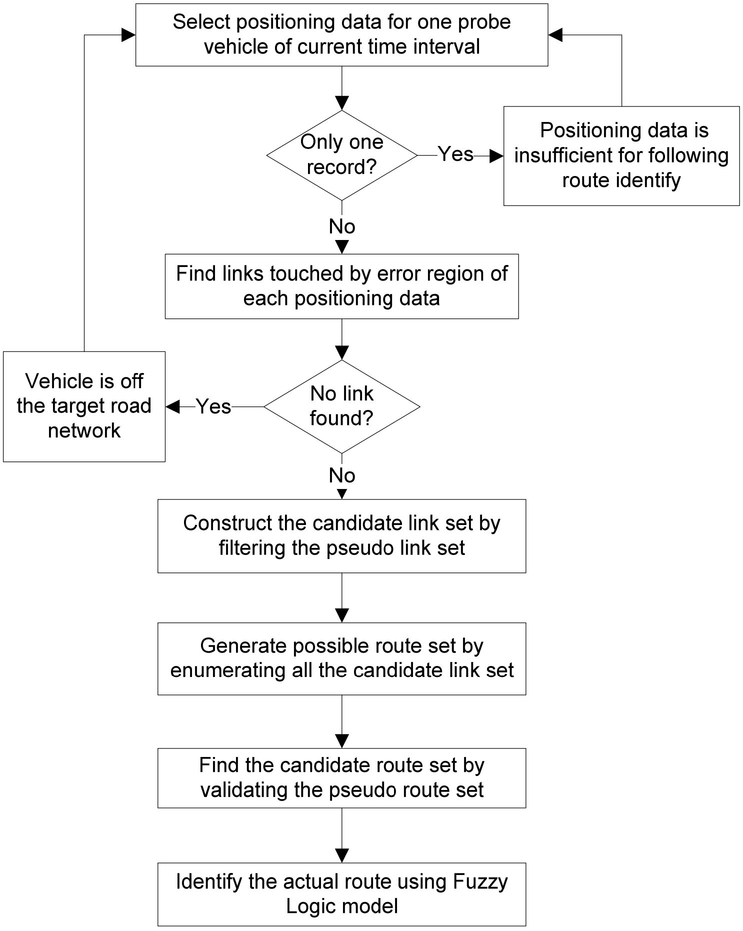


FIGURE 2 Flow chart of the trajectory tracking procedure (TTP).

Due to the GPS horizontal error for moving vehicles, the positioning points may fall in a region around the true position, which is usually named as error region or confidence region. Many methods are available for calculating the error region. Variance-covariance information associated with GPS receiver outputs is often used to define an error ellipse. For more details about the error eclipse calculation we refer readers to Ochieng W.Y. et al(2004)*(14)*. To simplify the process and reduce the computation complexity, the error region is defined as a rectangle in this study. The size of the error rectangle is based on the error analyzing of the GPS positioning data. All the links within the error region of positioning data *pl* are taken as the pseudo candidate links for pl, $\{L^{t,pl}_{T,i}$, *pl* = 0, 1, 2 …}. If there is no links in the error region, it means the vehicle is off the known freeway network at that time.

Then, a filtering process needs to be carried out to eliminate invalid candidate links from $L^{t,pl}_{T,i}$. The input variables for this process include minimum distance between a positioning point and a roadway link, similarity of roadway azimuth and vehicle heading. The technique to find the minimum distance is to project the positioning point to the target link and calculate the perpendicular distance (*PD*) *(2, 4)*. Because an link is composed of a finite sequence of points, we can also think that an link is composed of “straight lines”(or “inner lines”). We calculate the PDs between a positioning point and each inner line of a link, and select the smallest value as the perpendicular distance between the point and the link. The technique used in this study to find the PD between a point and a straight line is represented by equation (1) and (2). Given a link defined by two end points $P_1(x_1, y_1)$ and P_2 (x_2, y_2), a GPS positioning data $P_3(x_3, y_3)$, the perpendicular point $P(x, y)$ and the perpendicular distance can be calculated using following equation:

$$P(x,y) = \begin{cases} x = x_1 + \mu\,(x_2 - x_1) \\ y = y_1 + \mu\,(y_2 - y_1) \end{cases} \tag{1}$$

Where:

$$\mu = \frac{(x_3 - x_1)(x_2 - x_1) + (y_3 - y_1)(y_2 - y_1)}{\|P_2 - P_1\|^2}$$

$$PD = ER \times \arcsin\left(\sqrt[2]{\left(\sin\left(\frac{a}{2}\right)\right)^2 + \cos(R_{x3}) \times \cos(R_x) \times \left(\sin\left(\frac{b}{2}\right)\right)^2}\right) \tag{2}$$

Where:

$R_{x3} = (x_3 \times \pi) / 180$

$R_{y3} = (y_3 \times \pi) / 180$

$R_x = (x \times \pi) / 180$

$R_y = (y \times \pi) / 180$

$a = R_{y3} - R_y$

$b = R_{x3} - R_x$

ER: The earth Radius (Mile)

The difference between GPS heading and roadway link bearing is used to describe the similarity of roadway azimuth θ and vehicle heading Ψ, which is named as heading error (*HE*). Given a link defined by two end points $P_1(x_1, y_1)$ and P_2 (x_2, y_2), θ and the *HE* can be calculated by following equations:

$$\theta = \begin{cases} \left(2.5 \times \pi - \arctan(\frac{\Delta y}{\Delta x})\right) \times \frac{180^{o}}{\pi},\ if\ \Delta x < 0\ and\ \Delta y \geq 0 \\ \left(0.5 \times \pi - \arctan\left(\frac{\Delta y}{\Delta x}\right)\right) \times \frac{180^{o}}{\pi},\ others \end{cases} \tag{3}$$

Where:

$\Delta x = x_2 - x_1$

$\Delta y = y_2 - y_1$

$$HE = \begin{cases} |\Psi - \theta|, & if\ |\Psi - \theta| \leq 180^{o} \\ 360^{o} - |\Psi - \theta|, & if\ |\Psi - \theta| > 180^{o} \end{cases} \tag{4}$$

Ochieng (2004)*(14)* pointed out that GPS heading is not reliable for vehicle speed lower than 3 m/s (7 mph). So two different schemas are developed to select the candidate link set $\{L_{T,i}^{t,cl}$, $cl = 0, 1, 2 \ldots\}$ base on the vehicle speed estimated by GPS. When GPS speed is equal or larger than 7 mph, for all links with $HE \leq 45°$ (45° was selected based on the trial run), links with top two shortest *PD*s are selected as candidate links. When GPS speed is less than 7 mph, links with the top four shortest *PD*s are selected as candidate links.

Given the obtained candidate link set $L_{T,i}^{t,cl}$ for each positioning point, the next step is conducted the modified N-shortest path analysis to generate the candidate routes. Here we assume that the probe vehicles always take the route with shortest travel time between two continues GPS positioning points. Because the time gap between two continues GPS positioning points usually within several minutes or even seconds, our assumption is reasonable. Based on this assumption, we design our own N-shortest path algorithm. The algorithm is based on Dijkstra algorithm, proposed by Edsger Dijkstra*(30)* in 1959.

First, all the candidate links for each positioning point are listed and then candidate link pairs are generated for each GPS point pair. Second, topological analysis is conducted to eliminate any unconnected link pairs based on shortest path searching. When shortest path searching cannot return a reasonable route (e.g. route with reasonable length given physical vehicle speed limitation) between a link pair, it is considered as un-connected. Then a candidate route set for vehicle i at time slot T, $\{R_{T,i}^{cr}$, $cr = 0, 1, 2 \ldots\}$ is generated by enumerating all possible combinations of routes found in the previous step and the routes are listed in ascending order of travel time. At last, a Fuzzy Inference System (FIS) based on Fuzzy Logic theory is used to find the actual route $\mathrm{R}_{\mathrm{T,i}}^{\mathrm{r}}$ among the candidate route set $\{R_{T,i}^{cr}$, $cr = 0, 1, 2 \ldots\}$. Details of the FIS procedure will be presented in the latter section.

Position Determination Procedure (PDP)

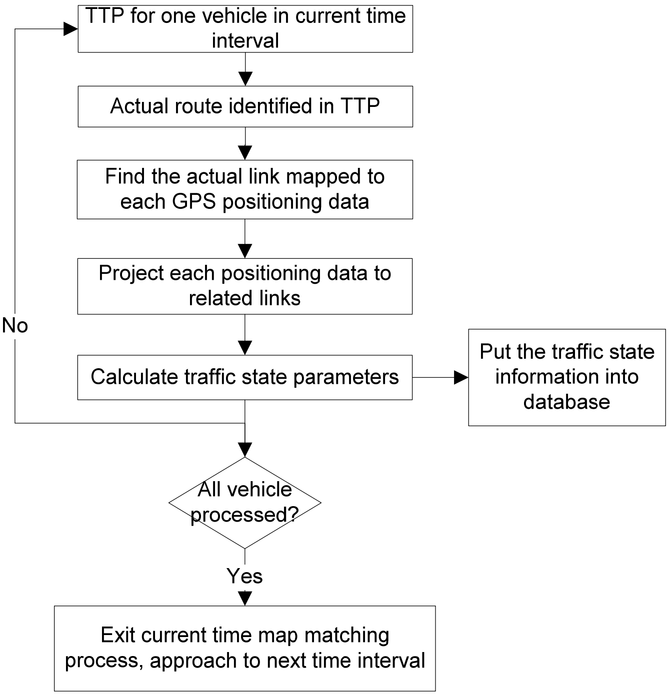


FIGURE 3 Diagrammatic representation of Position Determination Procedure (PDP).

After we found the actual route $R_{T,i}^{r}$ for probe vehicle i in time interval T in the previous TTP stage, we still need to know the probe vehicle's accurate physical location on the route at each timestamp of the positioning points. FIGURE 3 illustrates the framework of PDP in this study. The physical location of a GPS point can be obtained by projecting the positioning points to its corresponding link in the selected route, which has been stored at the TTP stage. The perpendicular position can be calculated by equation 1.

Fuzzy Logic Model

Because of the error of GPS positioning, the digital map error and the complexity of the road network, it is difficult to identify the true route. Usually the result we get from map matching procedure is the likelihood of a vehicle on a certain route. Consequently, techniques for dealing with qualitative terms such as likeliness are important in map matching algorithms *(19, 28, 31)*.

Fuzzy logic is an effective way to deal with qualitative terms linguistic vagueness and human intervention (Zhao, 1997)*(32)*. A Fuzzy Inference System (FIS) is a robust approach for building complex and nonlinear relationship between input and output data using fuzzy logic theory. In this study, a zero-order Sugeno-type FIS is used. The implementation of the FIS can be described as the following three steps: 1) fuzzification of the input and output, 2) formulation of the fuzzy rules, and 3) defuzzification of the output.

The state input variables in this FIS are: 1) the average speed of the probe vehicle in one study period (or average speed of the trajectory), ACS (MPH), 2) average heading error of the trajectory, AHE (AHE $= \frac{\sum_{\mathrm{i}=1}^{\mathrm{n}} HE_{\mathrm{i}}}{n}$)(Degree), 3) average distance between the candidate route and the vehicle trajectory, APD (Mile), 4) travel time of the candidate route, TTR. The fuzzy subsets associated with AVT are "zero", "low" and "high". For AHE and APD, the fuzzy subsets are "small" and "large". For TTR, they are "shortest" and "not shortest". Z-shaped, S-shaped, Bell-shaped and sharp-edged membership functions are chosen in the fuzzification process. FIGURE 4 shows the fuzzification result of the four state input variables. The output of this FIS is the likelihood of matching the probe vehicle trajectory to a candidate route, denoted as Z. The fuzzy subset associate with Z is low (Z_1) = 10, average (Z_2) = 50, high (Z_3) = 100.

The next step is to construct the fuzzy rules. Zhao et al,. (1997)*(32)* derived eight fuzzy rules for the case in which the positioning data came from a DR sensor. Quddus et al., (2006)*(19)* refined them into six rules based on their engineering knowledge. In this research we formulate six fuzzy rules based on the number of the state variables. According to the work of Quddus et al., (2006)*(19)*, Quddus et al., (2003)*(10)* and Greenfeld (2002)*(7)*, the AHE should be given more weight than APD. And to surmount the effect of the network complexity, especially the ramps and frontage roads, this study suggests to assign more weight to TTR too. Therefore, a higher weight is given to the rules associated with AHE and TTR.

- If (ACS is high) and (AHE is small) then ($Z = Z_2$) (3)
- If (ACS is high) and (AHE is large) then ($Z = Z_1$) (1)
- If (AHE is small) and (APD is low) then ($Z =Z_3$) (1)
- If (AHE is large) and (APD is high) then ($Z =Z_1$) (1)
- If (TTR is shortest) then ($Z =Z_3$)(2)
- If (TTR is not the shortest) the ($Z =Z_1$)(1)

To defuzzify the output, the min (minimum) method is used to calculate the "degree of applicability" (ωi) of each fuzzy rule. This FIS is applied to each candidate route in the candidate route set. The route with the highest likelihood is chosen as the actual route.

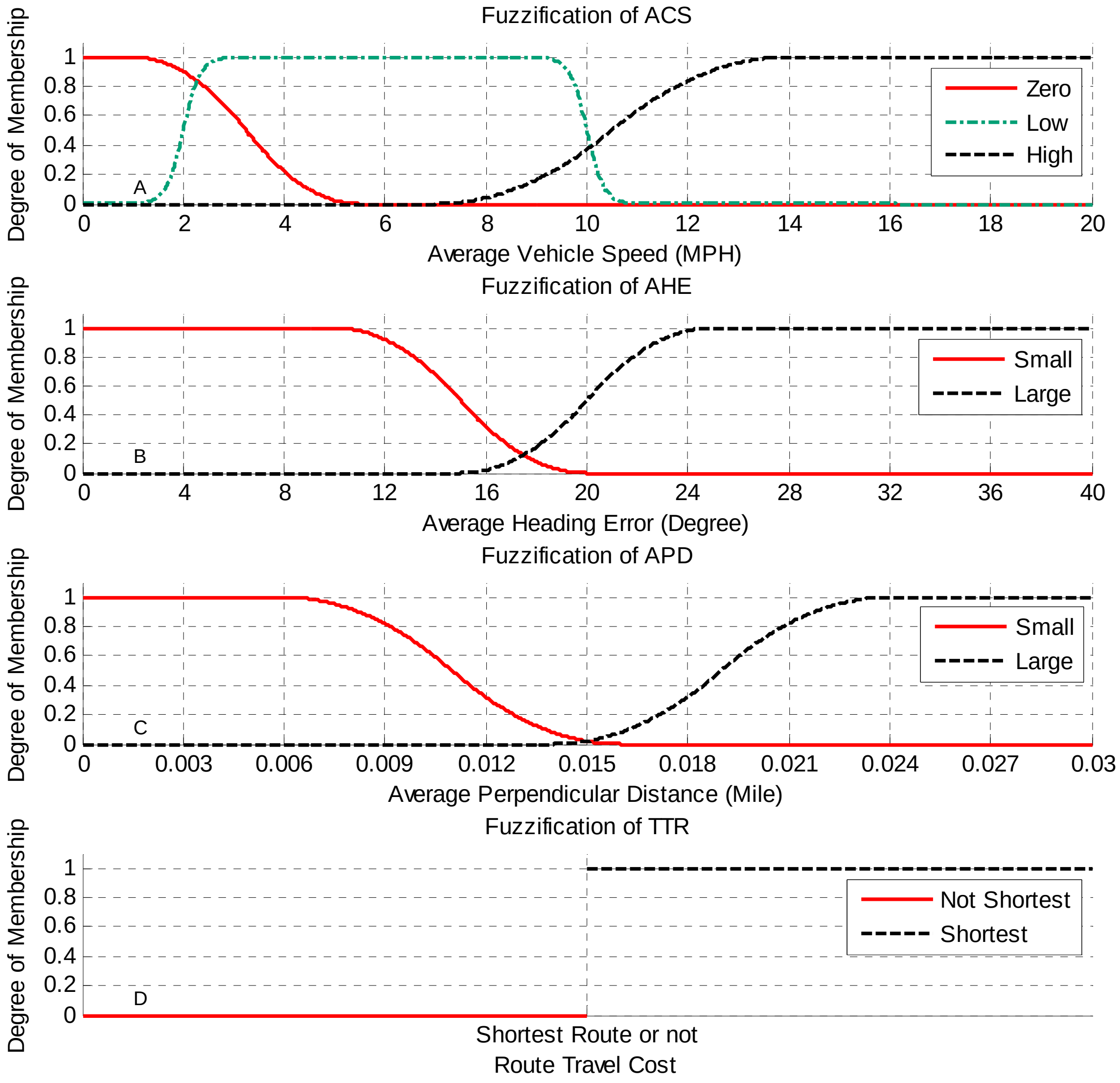


FIGURE 4 Fuzzification of input variables for FIS.

IMPLEMENTATION AND EVALUATION

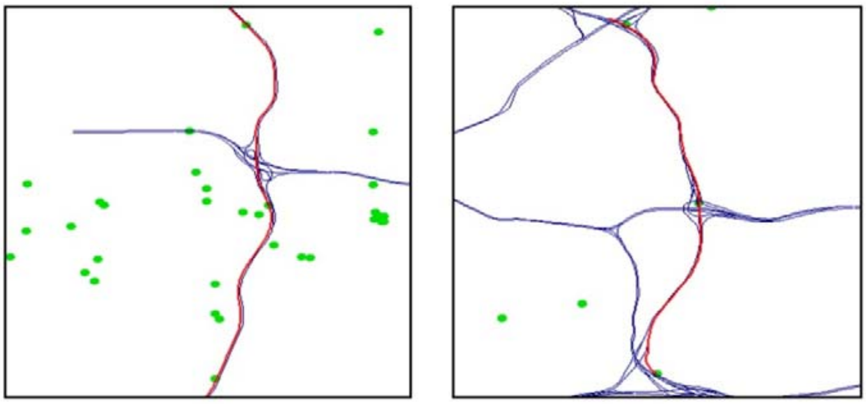
Field Test Implementation

It is important to evaluate the performance of this map matching algorithm in real-world applications. Real-life GPS positioning data and digital freeway network was collected in this field test. The network used is in the urban area of Los Angeles, CA consisting of more than 10,000 freeway links. The average length of the digital road links is 300 meters and there are about five shape points on each link. GPS data used in this test were collected from fleet management system on Wednesday, April 8, 2009. Also, the loop detector data for this freeway network was also collected and correlated with the GPS speed detection results for evaluation.

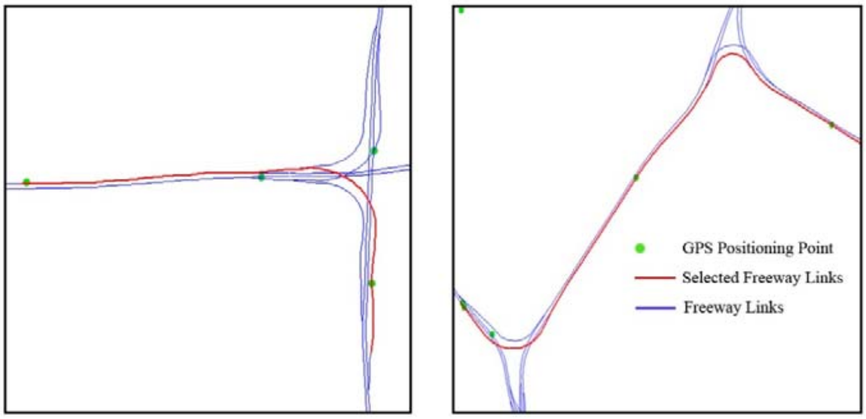
The algorithm was coded in Java version 6 by Eclipse IDE 3.4.1. The database used to store and computation data was Oracle Enterprise 10.2.0.The program ran on a notebook PC equipped with Intel Core 2 processor clocked at 1.8 GHz and 3 GB of RAM. The analyzing time interval *T* chosen by current traffic information industry usually is five or fifteen minutes. In this field test we selected five minutes, which means the map matching algorithm runs recursively for every five minutes. The error rectangles' size was 25 meters by 25 meters. According to the implementation results, the proposed algorithm can process about 170 GPS positioning points per seconds on the laptop described above.

Result Evaluation

Two methods were used to evaluate the performance of the algorithm. First, a manual check was carried out to check the accuracy of the map matching. This method was labor intense, and was only proper for checking results at locations with complex network geometries. Second, traffic speed information collected by loop detectors in Los Angeles was used as ground truth data to evaluate the traffic speed calculated by the GPS probe system using the proposed algorithm as the map matching module. In this way, map matching results are converted to a new data set that can be compared easily with ground truth data.



Vehicle ID: 45276182 Vehicle ID: 45275484



Vehicle ID: 45275272 Vehicle ID: 3151476

FIGURE 5 Performance of the proposed map matching algorithm in complex situations.

Some “hot spots” such as interchange and frontage roads were checked using the first method. Generally, if the algorithm could handle those complex situations, it implies that the algorithm can perform correctly for other part of the freeway network. The four examples in FIGURE 5 illustrates that even when some probe vehicles only returned two or three positioning points in one analyzing time interval *T*, the algorithm can find the actual routes for each probe vehicle accurately. On the other hand, we selected a specific area and carried out limited manual checking and found the percentage of correctly matched links was over 98%.

There were two approaches to accomplish the second method. The first approach was macroscopic level checking by comparing the average traffic speed of the freeway network collected independently by GPS probes and loop detectors. The second approach was link-based microscopic analysis, for those links which had both GPS probes and loop detectors at same time interval. Using the speeds from loop detectors as ground truth, the speed estimated by this algorithm was examined.

FIGURE 6 illustrates the analysis result from macroscopic level. There are two possible reasons for the obvious systematic difference in FIGURE 6. First, according to May(1990)*(33)*, the space-mean speed (SMS) measured by GPS probes should always be smaller than the time-mean speed (TMS) collected by loop detectors. Second, both GPS probes and loop detectors have sample bias problem. The GPS probes in this study were primarily freight vehicles, whose driver behavior is usually conservative. Dowling(1996) *(34)* addressed that the loop detectors will see more high-speed vehicles than slower vehicles passing a given point during a fixed time period. So the problem of sample bias determines that the speed measure by GPS probes should be lower than the speed collected by loop detectors. Despite the reasonable systematic difference of speed value, the perfectly matched trend of the two speed curves represents that the results of the proposed algorithm is reliable.

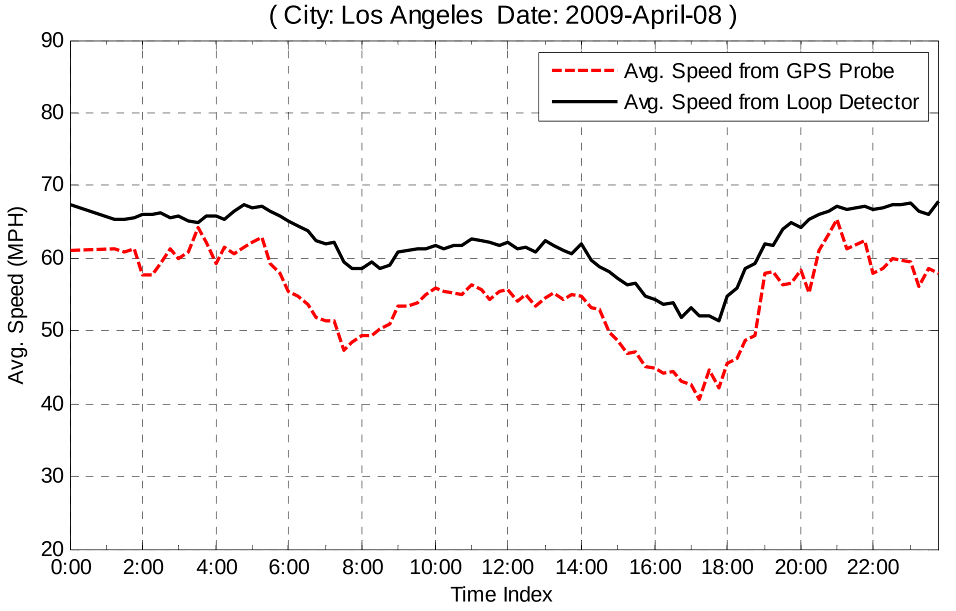


FIGURE 6 Freeway average speed comparison between GPS probes and loop detectors.

FIGURE 7 shows the comparison results from microscopic level. As indicated in FIGURE 7-A and 7-B, the speed detected by GPS probes and loop detectors are quite consistent for those links with heavy traffic congestion. FIGURE 7-C and 7-D illustrates that the speed from GPS probes will be consistently lower than speed from loop detectors for those links without any major congestion. Sometimes the GPS positioning data sets are not as good as expected due to incidents and GPS signal issues. This will produce outlier in speed estimation using GPS probe data sets. As shown in FIGURE 7-E and 7-F, the outliers need to be filtered or smoothed for traffic state monitoring purpose. The loop detectors could be unreliable when malfunctioning or in lack of calibration or maintenance. The two loop detectors in FIGURE 7-G and 7-H obviously have some problem and are reporting the default speed during the test day while the speed from GPS probes is more reasonable in this situation.

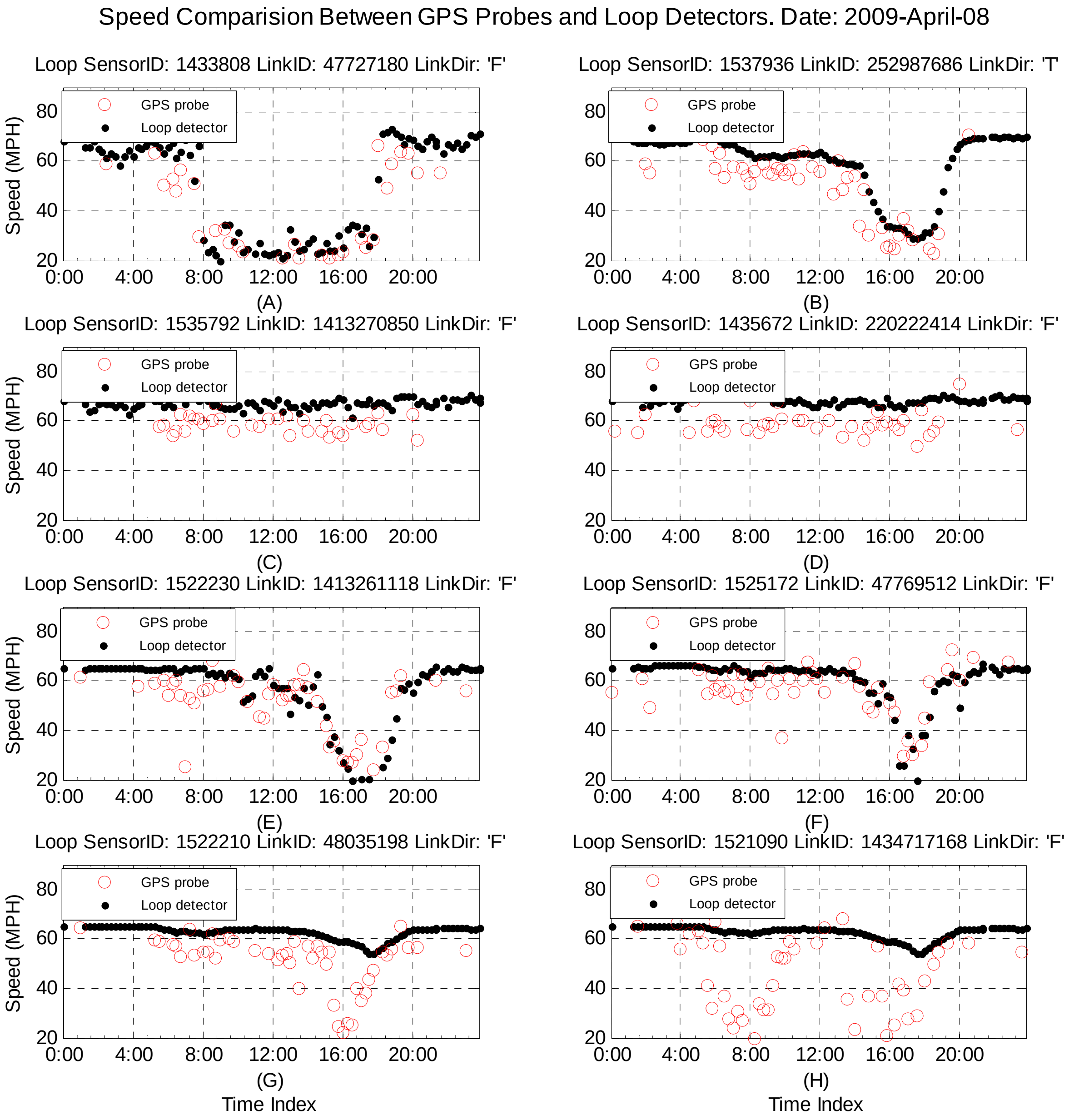


FIGURE 7 Map matching results checking using independent observation.

CONCLUSION

This paper proposed a new integrated map matching algorithm for large-scale network-level traffic monitoring based on GPS probe data. In previous literatures, only a few map matching algorithms have been proposed for this application. Such application requires the map matching algorithm to be able to 1) simultaneously map a series of GPS positioning points with temporal difference ranging from a few seconds to a few minutes, 2) handle complex network geometries with accuracy and robustness, and 3) process large amount of data with reasonable computational cost. In order to fulfill these requirements, the proposed algorithm uses several novel methods. The first one is the modified N-Shortest path algorithm specially design to find reasonable candidate routes between two GPS points that may be a few minutes apart. To solve the uncertainty and signal noise problems commonly experienced by GPS data, the fuzzy logic model is applied to enhance the existing probabilistic map matching algorithms for the selection of actual route from candidate routes and the determination of actual location of a GPS positioning point. To reach a low computational cost, several simplification techniques such as interval-by-interval processing, point error region simplification and so on are used.

The algorithm is evaluated based on two strategies. The first one is the traditional case-based visual inspection of mapping results. The second one is speed comparison based strategy. This new evaluation strategy was first introduced in this paper for map matching algorithms serving for traffic monitoring purpose. First, link speeds are calculated using GPS probe system that uses the proposed algorithm as the map matching module. Then, the detected speed results are compared with loop detector speed collected at corresponding roadway links. In this way, map matching results are converted to a new data set that can be easily compared with actual ground truth. The evaluation results following both strategies illustrate the effectiveness and robustness of the algorithms.

Meanwhile, limitations are also found. First, a critical problem caused by “interval-by-interval” processing is the unnecessary division of vehicle trajectories that pass across the interval boundaries. This can cause the loss of route connectivity information resulting in un-reasonable routes. Second, the low penetration rate can cause issue for map matching. The algorithm can only deal with GPS pair that has reasonable time differences, for example, within 5 or 15 minutes. GPS pairs with long time differences are considered as invalid. Third, the fuzzification process requires expert knowledge after analyzing the results from several trial runs, which may be time-consuming. However, with the help of the second evaluation strategy, the calibration time can be seriously reduced.

Future study for the topic includes several directions. First, more comprehensive statistical analysis based on the second strategy. Second, conduct more visual inspection under the first strategy so that the sensitivity of the proposed algorithm with respect to different network conditions can be evaluated. Third, we only conducted mapping on freeway networks, however, the algorithm has the potential to be extended to arterials since they can all be covered by GPS probe vehicles.

REFERENCES

1. Quddus, M.A., W.Y. Ochieng, and R.B. Noland, Current map-matching algorithms for transport applications: State-of-the art and future research directions. *Transportation Research Part C: Emerging Technologies*, Vol. 15, 2007, pp. 312-328.
2. Bernstein, D. and A. Kornhauser, *An introduction to map-matching for personal navigation assistants*. Princeton University; New Jersey TIDE Center. 1996.
3. Kim, W., G.-I. Jee, and J. Lee, Efficient use of digital road map in various positioning for ITS. *IEEE Symposium on Position Location and Navigation*, 2000.
4. White, C.E., D. Bernstein, and A.L. Kornhauser, Some map matching algorithms for personal navigation assistants. *Transportation Research Part C: Emerging Technologies*, Vol. 8, 2000, pp. 91-108.
5. Joshi, R.R. A new approach to map matching for in-vehicle navigation systems: the rotational variation metric. in *IEEE Intelligent Transportation Systems*. 2001.
6. Pyo, J.-S., D.-H. Shin, and T.-K. Sung. Development of a map-matching method using the multiple hypothesis technique. in *IEEE Intelligent Transportation Systems Conference.* 2001.
7. Greenfeld, J.S. Matching GPS Observations to Locations on a digital Map. in *Transportation Research Board 81st Annual Meeting*. Washington, D.C. 2002.
8. Gustafsson, F., et al., Particle filters for positioning,navigation, and tracking. *IEEE Transactions on Signal Processing*, Vol. 50, 2002, pp. 425-435.
9. Cui, Y. and S.S. Ge, Autonomous vehicle positioning with GPS in urban canyon environments. *IEEE Transactions on Robotics and Automation*, Vol. 19, No. 1, 2003, pp. 15-25.
10. Quddus, M.A., et al., A general map-matching algorithm for transport telematics applications. *GPS Solutions*, Vol. 7, No. 3, 2003, pp. 157-167.
11. Yang, D., B. Cai, and Y. Yuan. An improved map-matching algorithm used in vehicle navigation system. in *IEEE Proceedings on Intelligent Transportation Systems*. 2003.
12. Huabei, Y. and O. Wolfson. A weight-based map matching method in moving objects databases. in *16th International Conference on Scientific and Statistical Database Management*. 2004.
13. Mengyin, F., L. Jie, and W. Meiling. A hybrid map matching algorithm based on fuzzy comprehensive judgment. in *The 7th International IEEE Conference on Intelligent Transportation Systems*. 2004.
14. Ochieng, W.Y., M. Quddus, and R.B. Noland, Map-matching in complex urban road networks. *Brazilian Journal of Cartography*, Vol. 55, No. 2, 2004, pp. 1-18.
15. Ochieng, W.Y., M. Quddus, and R.B. Noland. Integrated positioning algorithms for transport telematics applications. in *ION GNSS 17th International Technical Meeting of the Satellite Division*. Long Beach, CA. 2004.
16. Winter, M. and G. Taylor. Modular Neural Networks for Map-Matched GPS Positioning. in *Fourth International Conference on Web Information Systems Engineering Workshops*: Institute of Electrical and Electronics Engineers. 2004.
17. Blazquez, C.A. and A.P. Vonderohe, Simple Map-Matching Algorithm Applied to Intelligent Winter Maintenance Vehicle Data. *Transportation Research Record: Journal of theTransportation Research Board*, Vol. 1935, Transportation Research Board of the National Academies, Washington, D.C., 2005, pp. 68-76.
18. Fontaine, M.D. and B.L. Smith, Probe-Based Traffic Monitoring Systems with Wireless Location Technology: An Investigation of the Relationship Between System Design and Effectiveness. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 1925, Transportation Research Board of the National Academies, Washington, D.C., 2005, pp. 3-11.

19. Quddus, M.A., R.B. Noland, and W.Y. Ochieng, A High Accuracy Fuzzy Logic Based Map Matching Algorithm for Road Transport. *Journal of Intelligent Transportation Systems*, Vol. 10, No. 3, 2006, pp. 103-115.
20. Hongchao, L., et al. Combined Kalman Filter and Global Positioning System Error Correction Approach for Enhanced Map Matching. in *Transportation Research Board 87th Annual Meeting*: Transportation Research Board. 2008.
21. Nassreddine, G., F. Abdallah, and T. Denoeux. Map matching algorithm using belief function theory. in *11th International Conference on Information Fusion*. 2008.
22. taghipour, S., M.R. Meybodi, and A. Taghipour. An Algorithm for Map Matching For Car Navigation System. in *3rd International Conference on Information and Communication Technologies: From Theory to Applications*. 2008.
23. Marchal, F., J. Hackney, and K.W. Axhausen, Efficient Map Matching of Large Global Positioning System Data Sets: Tests on Speed-Monitoring Experiment in Zürich. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 1935, Transportation Research Board of the National Academies, Washington, D.C., 2005, pp. 93-100.
24. Pan, C., et al., Data Collection Based on Global Positioning System for Travel Time and Delay for Arterial Roadway Network. *Transportation Research Record: Journal of theTransportation Research Board*, Vol. 2024, Transportation Research Board of the National Academies, Washington, D.C., 2007, pp. 35-43.
25. Sinn, K. and K. Jong-Hwan. Q-factor map matching method using adaptive fuzzy network. in *IEEE International Fuzzy Systems Conference*. 1999.
26. Quddus, M.A., W.Y. Ochieng, and R.B. Noland, Integrity of Map-Matching Algorithms. *Transportation Research Part C: Emerging Technologies*, Vol. 14, No. 4, 2006, pp. 283-302.
27. Syed, S. and E. Cannon, Map-Aided GPS Navigation: Linking Vehicles and Maps to Support Location-Based Services. *GPS World*, Vol. 16, 2005, pp. 39-44.
28. Su, H., J. Tang, and C. Hou. A Integrated Map Matching Algorithm Based on Fuzzy Theory for Vehicle Navigation System. in *International Conference on Computational Intelligence and Security*. 2006.
29. USDOD, *Global Positioning System Standard Positioning Service Performance Standard*, 2001
30. Dijkstra, E.W., A note on two problems in connexion with graphs. *Numerische Mathematik*, Vol. 1, 1959, pp. 269-271.
31. Zhang, Y. and Y. Gao. A Fuzzy Logic Map Matching Algorithm. in *Fifth International Conference on Fuzzy Systems and Knowledge Discovery*. 2008.
32. Zhao, Y., *Vehicle Location and Navigation Systems*. Artech House Publishers, Norwood, MA. 1997
33. May, A.D., *Traffic Flow Fundamentals*. Prentice Hall. 1990
34. Dowling, R.G. and W. Cheng, Evaluation of Speed Measurement and Prediction Techniques for Signalized Arterials. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 1564, Transportation Research Board of the National Academies, Washington, D.C., 1996, pp. 20-29.