Map Matching of GPS Data with Bayesian Belief Networks

Tao FENG [a], Harry J.P. TIMMERMANS [b]

[a,b] *Department of the Built Environment, Eindhoven University of Technology, The Netherlands*
[a] *E-mail: ftaocn@hotmail.com*
[b] *E-mail: h.j.p.timmermans@tue.nl*

Abstract: This paper proposes a map matching algorithm using Bayesian belief network for GPS traces to generate the spatial-temporal information of individuals. The algorithm incorporates the road network topology, distance from trace nodes to road segments, the angle between two lines, direction difference, accuracy of measured GPS log point, and position of roads. The GPS data collected in the Eindhoven region, The Netherlands, was used to examine the performance of this algorithm. Results based on a small sample show that the algorithm has a good performance in both processing efficiency and prediction accuracy of correctly identified instances. Even with a small sample, the overall prediction accuracy reaches 87.02%.

Keywords: Algorithm, Connectivity, Distance, Network

1. INTRODUCTION

Map matching algorithm has been an important research topic in recent years in transportation. It has been used in general with a purpose to match some geographical locations or points with the existing network data. For vehicle-based traffic research, it can provide microscopic spatial-temporal information for a specific vehicle. This would be beneficial to the requirements of many research topics such as route choice simulation, emission and energy consumption analysis, demand forecast, travel behavior, etc. Especially, with the increasing applications of new technologies such as GPS, Bluetooth, smart phone in data survey, the requirement of an efficient and valid map matching algorithm to identify the exact locations of the trace points on road networks becomes extremely important.

Basically, to identify whether a node belongs to a specific road segment (link) is not straightforward because of the difference between the spatial locations of the node and the link, road network topology and the accuracy of the geographical measurement. This can be more complicated if a node is located in an area surrounded by high density buildings where the strength of the satellite signals heavily influences the accuracy of the location measurement. In addition the variety in some special cases, like U-turn, passing through a round-about, lane changing, also increases the complexity to design map matching algorithms. A well-performed model should not only handle these main research issues, but provide sufficient robustness and uncertainty measurement in the model to be applicable in different contexts.

Due to the fact that the GPS data fluctuate according to the contextual information, like the weather, urban density, sensitivity of the GPS sensor, etc., the traces measured in different time slot using a same device can result into different levels of accuracy. Therefore, it seems to be important that the designed algorithm incorporates a learning function into the model in the sense that the estimated parameters could be adjusted through an intelligent learning

procedure, which will then increase the overall accuracy of the map matching results with more data coming in.

A common approach of map matching which has been used empirically is by means of the spatial analysis functions provided by geographical information systems (GIS) tools. The buffering area along with the geographical objects, i.e. lines or nodes, are created first, and used to match other geographical data where the various criteria of overlapping filters apply. Such a method is popularly applied because of the convenience to implement, but it needs to set the threshold of the searching radius which can be different according to the accuracy of the measurement. For example, Du and Aultmann-Hall (2007) and Arentze et al. (2012) have used a 10 meter as a threshold value to create the buffer and match the road network with GPS points. Although this threshold value has empirically shown the acceptable accuracy, the method cannot meet the accuracy requirement of map matching for personal traces. In particular, the variation of the coordinates leads to a dilemma that some nodes cannot be recognized as belonging to any roads if the searching radius is not big enough, while a too big radius will on the other hand result into unrealistic matching results.

A prototype of map matching algorithms was developed by using the distance from node to node, the distance from nodes to lines and lines to lines (Bernstein and Kornhauser, 1996). These algorithms are consistent in the sense the distance was considered as the only decision variable. This can be problematic in some applications because of the ignorance of other important variables, like the connectivity of road segments. Although such algorithms have shown a good performance regarding the processing speed, the matching accuracy is difficult to ensure.

In recent, research appeared to incorporate more variables into the map matching algorithms, like the angle between the directions of two lines, the connectivity of network topology and accuracy of the GPS measurements (Quddus et al., 2006; Ren, 2012; Michel et al., 2013). From a methodological point of view, existing map matching algorithms vary from the ad hoc rules to these advanced learning-based algorithms, i.e. fuzzy set, Hidden-Markov algorithm, probability hypothesis, etc. For instance, Manifest (1999) proposed the multiple hypothesis probability (MHP) algorithms for map matching. The method was further adopted and extended to the application of matching with GPS data (Schuessler and Axhausen, 2009). Quddus et al. (2006) introduced the fuzzy logic theory to map matching with road networks. Some of the important variables, like the search space based on the error ellipse derived from the error variances, the perpendicular distance from a position fix to the link, the bearing of the link, and the direction of the vehicle, are incorporated. Six rules were created for the fuzzy inference. The model was shown as to be more accuracy than previous models. However, it needs some expert knowledge to determine the fuzzy set.

Ren (2012) has checked on the existing map matching issue and proposed advanced algorithms for pedestrian and/or wheelchair navigation services. A Hidden-Markov model and a so-called multi-sensor approach which incorporates the accelerometer into map matching have been developed. In most recent, Michel et al. (2013) proposed a probabilistic algorithm which can be very useful for smart phone data. The model showed a quite good efficiency in the accuracy and provided an opportunity to measure the uncertainty of the candidate road sets. However, the model needs some simplified assumptions in real applications. Detailed literature reviews on the main existing map matching algorithms have been given by White et al. (2000) and Quddus et al. (2007).

Unlike the previous methods, the advanced algorithms can provide much potential values to handle the complexity of map matching problems. A well-performed algorithm should be flexible enough to measure the uncertainty of matched roads, and be capable to incorporate the major influential factors to ensure the prediction accuracy. Due to the fact that

few of the existing research investigated all these specifics in all, there is still a need to further improve the performance of the algorithms. Considering the increasing applications using Bayesian Belief Network (BBN) technique in different research fields and the superiority of BBN relative to other algorithms (Rudloff and Ray, 2010; Feng and Timmermans, 2013), it seems to be necessary to examine the feasibility of this algorithm in map matching issue. Therefore, in this paper, we will develop a map matching algorithm which incorporates the inference process using BBN model. The model incorporates most of the major influential factors, including the distance to road, the connectivity between two road segments, the direction difference of two line objects, the accuracy of GPS measurement and the direction difference between two adjacent links. The method is evaluated using the GPS data collected in the Eindhoven region, The Netherlands.

The remainder of the paper is organized as follows: Section 2 presents the algorithm of Bayesian belief network technology by representing the main influential factors and the modeling structure. Section 3 briefly describes the GPS data we used in this paper and the implementation issue. Section 4 presents the results. The paper will be summarized and concluded in Section 5.

2. THE ALGORITHM

To identify the road segment matched with a GPS point, it is necessary to incorporate the influential factors. In the following discussion, we will first illustrate the main influential factors we adopted, followed by presenting the proposed algorithm.

2.1 Control Variables

- **Distance to road**

The distance between two geographical objects represents the correlation between spatial locations. The road segment which is mostly closed to the referenced node has a high probability to be the correct road. However, the distances from a GPS point to a specific road segment vary in terms of the spatial correlation between the two geographical objects. As shown in Figure 1, there could be three types of distances between a node to a line, which are the perpendicular distance (d_0) and the distances from the node to the end nodes (d_1 and d_2) of the link.
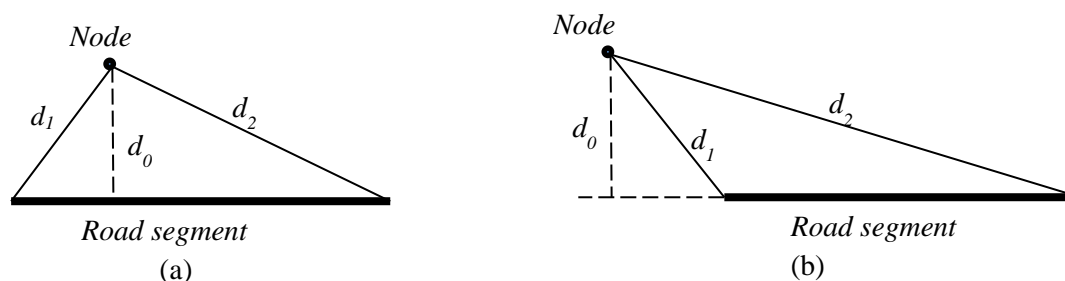

Figure 1 Distance between the a node to a road segment

In most of the cases, the perpendicular distance was considered as an efficient measurement. Other distances might be used to calculate angle information of the virtual triangular to provide further location information. In our case, we take the closest distance

between the node and the road. In another words, it is the perpendicular distance (d_0) as in the case of Figure 1a and d_1 as in the case of Figure 1b. In this way, every road segment in the candidate set can be differentiated through the shortest distance from a node to a road.

- **Direction difference**

Despite that the distance measures the reference information for a single node it does not say anything about the sequence information. The integral consideration of two or more adjacent nodes can provide information in both time and direction dimensions. Basically, the direction between the line connecting two adjacent nodes should be consistent in some extent with the direction of the matched link. Moreover, it can also benefit to the identification of whether there happens a turning action which mostly indicates a cross section closed to the node (Figure 2a).
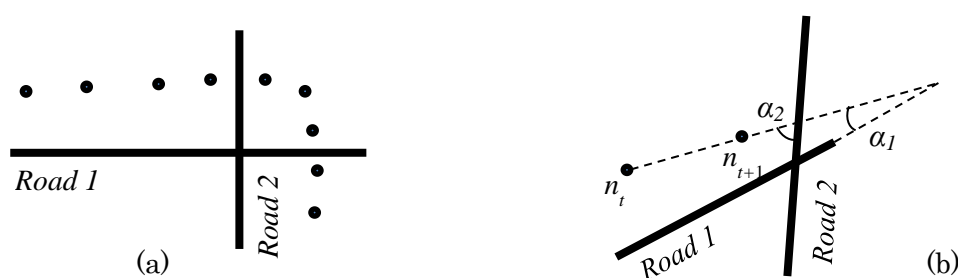

(a)   (b)

Figure 2 Spatial difference of directions between the trace line and road segments

To further identify the correlation between the line of two nodes and the road segment, we use the angle information by incorporating the direction of the lines, as shown in Figure 2b. In order to determine the degree of the angle consistently between a location and multiple road candidates, the direction of the road and two nodes are pre-set in advance. As shown in Figure 2b, to ensure that Road 1 and the link connecting the two nodes has to be in the same direction to confirm the angle of α rather than $180 - \alpha$. For the sake of simplicity, in practice, we restrict the angle as to be less than 90 degree.

- **Connectivity**

Apart from the distance and direction information, one of the most important factors is the connectivity which means the topology information of the road and cross sections in the geographical representation. This variable is quite important to filter out the un-reasonable roads and increase the searching efficiency. In practice, only these roads which are connected at least at one node with the previously confirmed road segment are treated as connected.

- **Other factors**

Since the measurement of GPS data are dependent on the signal strength, sensitivity of the sensor and the spatial context of the travel environment. In principle, the more number of satellites the GPS logger received, the less noise in the measurements. In addition, the difference in the complexity of traveling environment results into different travel patterns in the sense that the travel speed in the high density city center and suburban area differ significantly. Therefore, it is also necessary to incorporate the velocity information and accuracy of the GPS measurements.

2.2 The BBN Model

The method we adopted here is the Bayesian Belief Network (BBN) which replaces ad hoc rules with a dynamic structure, leading to improved classification if consistent evidence is obtained over time from more samples (more traces). Bayesian belief network (also called Bayes net) is a graphical representation of the conditional probability and causality relationships between variables. The model is described qualitatively by directed acyclic graphs where nodes and edges represent variables and the dependencies between variables. The nodes where the edge originates and ends are called the parent and the child, respectively. Bayesian belief networks allow for probabilistic inference to be performed, indicating that the probability of each value of a node can be computed when the values of the other variables are known.

The nodes that can be reached from other nodes are called descendent. In Bayesian network, each variable is independent of its non-descendent given the state of its parents. Since the independence among the variables are clearly defined, not all joint probabilities in the Bayesian system need to be calculated, which provides an efficient way to compute the posterior probabilities. Suppose the set of variables in a BBN is $\{A_1, A_2, \ldots, A_n\}$ and that $parents(A_i)$ denotes the set of parents of the node A_i in the BBN. Then the joint probability distribution for $\{A_1, A_2, \ldots, A_n\}$ can be calculated from the product of individual probabilities of the nodes:

$$P(A_i, \ldots, A_n) = \prod_{i=1}^{n} P(A_i | parents(A_i)) \tag{1}$$

In our case, the Bayesian belief network represents the multiple relationships between different spatial, temporal and other factors, including errors in the technology itself (input), and the facet of the candidate road segment that we wish to identify (output). We use a Bayesian belief network to impute automatically the probability for each road segment in a filtered road set.

Figure 3 shows the network structure that we use to infer the matched road segment with respect to GPS traces. A candidate road segment is treated as a function of the states of the variables included in the BBN.
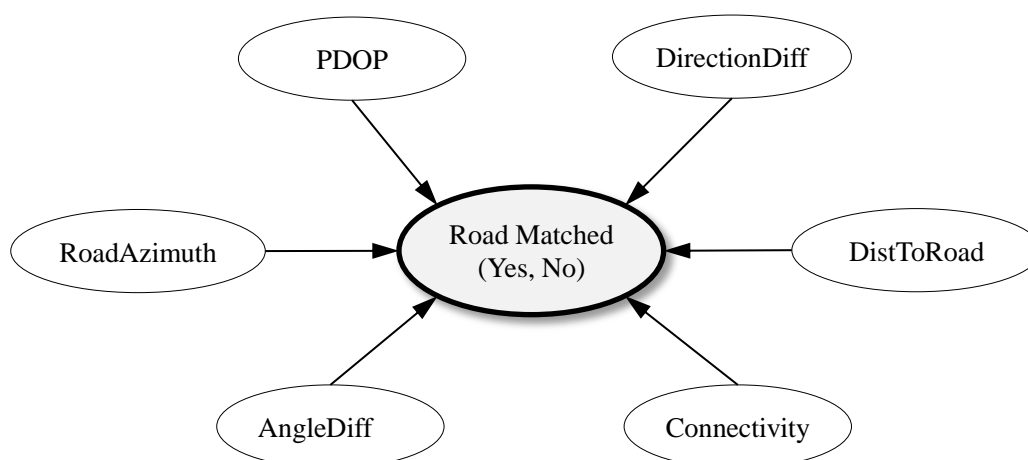

Figure 3 Model structure for the inference of map matching

To what extent a road segment is matched with a location may be partly determined by the information of the previously matched data. For example, traveling along a straight route will in general have a pattern that the directions of the adjacent links connecting two GPS log points are similar. While turning across a road section may result into significant difference between the directions. Therefore, we consider three variables as the sequence information, the connectivity (Connectivity), the angle difference between the directions of two adjacent links (AngleDiff), and the direction difference between the link connecting two nodes and the matched road segment (DirectionDiff).

Here, the connectivity indicates whether two adjacent road segments are connected each other. For example, to impute the probability of road segment i, we take the previously matched road segment i-1 as a reference, and check whether the current road i connects with the road i-1. This means we keep the information of the previous matched road segment temporally to identify the potential road segment for the current location. In case that two adjacent locations identified a same road segment i, the connectivity is set as true.

The DirectionDiff can be calculated as mentioned previously. The angle difference (AngleDiff) is the absolute value of the difference between two adjacent directions of the virtual links, as mentioned above. In case of the starting node, the direction is set to 0. Therefore, the angle difference between the second link and the previous link equals to the value of the direction with respect to the second GPS log point.

In the model, we include two variables for the accuracy of the measurements, the Position Dilution of Precision (PDOP) and the number of satellites (NSATS). The variable PDOP is a measure of overall uncertainty of a GPS position, represents the quality of GPS signals. A PDOP value of 1 indicates a good satellite configuration and high-quality data; conversely, PDOP values above 8 are considered poor. The quality of the data decreases as the PDOP value increases.

The variable (DistToRoad) is the perpendicular distance from a node to a road segment as presented above (Figure 1). The variable of RoadAzimuth measures the angle of the matched line object in a spherical coordinate system. Assume that a vector from an origin to a point of interest is projected perpendicularly onto a reference plane. Then the angle between the projected vector and a reference vector on the reference plane is called the azimuth. Therefore, the azimuth information indicates position of a road, which may provide useful input in combination with other location variables.

The output variable is whether a road segment is the matched road. In this case, it has two levels, yes or no. For each candidate road segment, we infer the probability through the conditional probabilities among input variables and the output variable. For a set of found options, the road segment which has the highest probability is taken as the rightly matched road.

3. GPS DATA AND IMPLEMENTATIONS

The GPS data used in this paper are collected in the urban area of Eindhoven, The Netherlands. One of the main purposes of this GPS data collection is to find valuable merits of peoples activity-travel trajectories in a long term. Every participant was required to join the survey for a period of three months. People carried the GPS logger, downloaded and uploaded their GPS data to the website. The data was then processed using the program of TraceAnnotator to generate the possible activity and travel data. Details about the TraceAnnotator can refer to the paper by Moisver et al. (2010). A web-based prompt recall procedure was used to ask people to validate their historical data in case of any incorrect data.

Additionally, respondents were also asked to provide some demography information and the locations where they visit frequently.

Individuals carried the GPS logger, named Bluetooth A++ Pro. The device has a good sensor embedded, which can receive well signals within trains. This capability decreases in some extent the noise induced by signals in urban area, however, the overall accuracy of the GPS data is in general rather acceptable especially for traveling data, like car for example. The GPS devices were configured to record data in every 3 seconds. The recorded information include: date, time, longitude, latitude, speed, distance, speed, accuracy of the measurement (like PDOP, PDOP, etc.), and number of satellites.
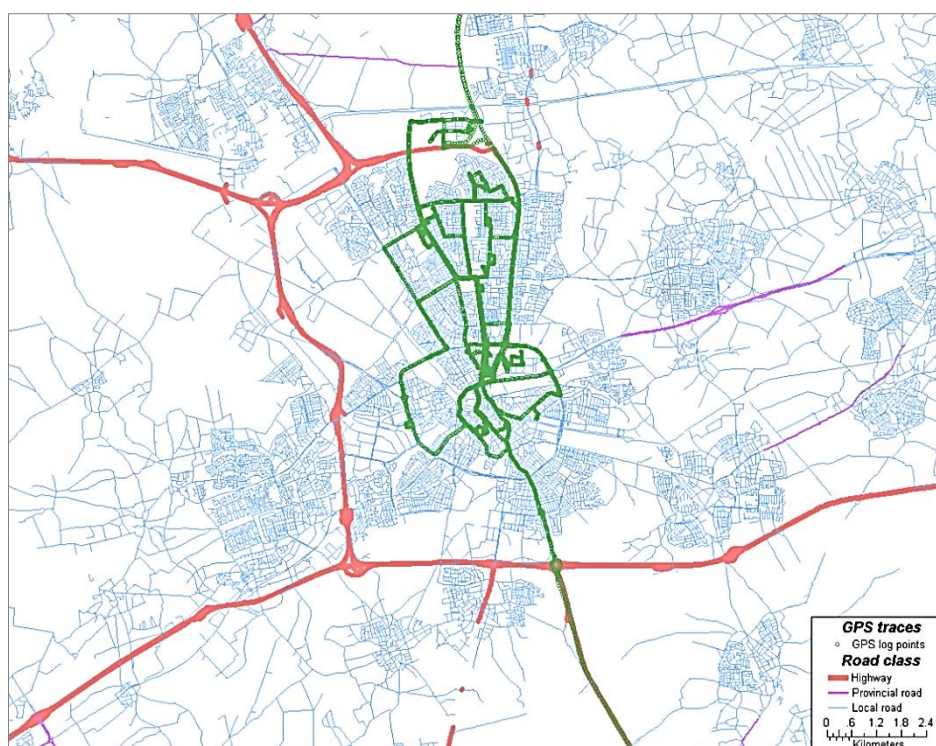

Figure 4 Distribution of the GPS log points on the road network of The Netherlands

Although the data collected in the Eindhoven region was targeted as a large scale of research projects where all available transportation modes and various activity types were included, in this paper, we are especially interested in the trips by car to test the efficiency of the proposed map matching algorithm. In addition, since the training of the BBN model needs some facts that can determine the conditional probabilities among the input variables and the road segment, we extract the trips by car which have detailed recorded diary.

The road network data includes all the road segments of the whole Netherlands. There are three categories of roads: the national road, the provincial road and the local road. Figure 4 shows the distribution of the GPS log points and road categories. As a consequence, a total of 10 trips, including 1227 GPS points are selected as a test sample.

In order to evaluate the efficiency of the proposed method, we labeled in the dataset the true road segments for every GPS log point using GIS. A filter was designed specifically to search the set of candidate road segments for a GPS coordinates. More specifically, a buffer with the given diameter (d_i) for a GPS point i was created based on the coordinates, and the road segments which touch with the buffer were taken into the choice set.

To determine a feasible value of search diameter will help improving the searching efficiency of the whole algorithms. Since the GPS traces are influenced by various indicators like number of satellites, weather, etc., the spatial pattern of the log points vary even for a same road segments. The empirical values of 2.5 and 11.4 have been used to classify the data quality of good and poor (Ren, 2012). As shown in Figure 5, the distances to the road from GPS log points of different days are different.

Because the searching process is memory intensive in that the searching time increases with the increase of the search diameter. Here, in order to increase the efficiency of the searching procedure, the search diameter was set dynamically according to the value of PDOP. We use 20 meters as the searching diameter if the PDOP has a value less than 3, and 50 meters otherwise.
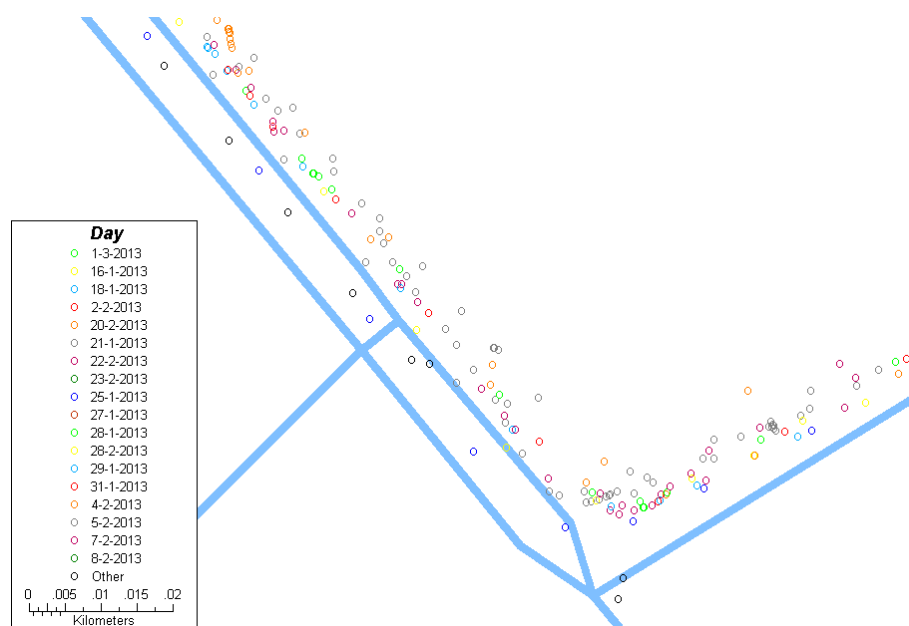

Figure 5 Distances from the GPS log points of different days to road segments

$$d_i = \begin{cases} 20, if\ PDOP \leq 3 \\ 50, if\ PDOP > 3 \end{cases} \tag{2}$$

In the process of BBN training, the filter was first implemented for every log point in the sample to generate a set of candidate road segments. Then the input variables in the BBN model were calculated for each road segment and the GPS coordinates. A new dataset where each row represents the pattern of matched or un-matched road segment and their relevant attribute variables was created. As presented above, we incorporated variables relevant to speed, connectivity, spatial distance to the specified road segment, direction and the accuracy of GPS log measurements.

The true sample data was prepared using the carefully recorded diary and the road network data within GIS. The locations which cannot match with any road segments were excluded. We divided the sample data into two datasets, one is for training purpose and the other is for test purpose. The datasets were selected randomly in terms of trips and road categories, by setting 75% and 25% as the training and test dataset, respectively.

Since the connectivity is considered as one of the determinant variables, it is important

to find the first road segment of the starting log point (origin). Because the effect of signal issue in the certain activity location, the exact coordinates is in general not accuracy enough. Moreover, a prompt recall cannot ensure the mentioned location as a truth. Therefore, we add some additional filters in assisting the identification of the origin. More specifically, we include the information of personal profiles and the validated location names in combination with the activity type. In concrete, if the personal profile is available in terms of the activity type, we picked the coordinates of their personal locations directly from the database of personal locations. Otherwise, the location names will be used for online geocoding to find a valid coordinates through Google service. Finally, if both the personal profiles and location names are missing or incomplete for some reasons, we fetch the original GPS traces according to the start time and/or end time of the activity. Finally, the coordinates of the origin is generated as to provide references of map matching for the consequent GPS log points. A detailed implementation procedure is depicted in Figure 6.

As shown in the flowchart, the map matching process starts from the origin and ends at the destination of a trip. The detection of real origin is necessary since the traces during activity episode fluctuate in general. Taking the first data record as the starting node can be sometimes problematic because the device needs some time to configure at the beginning of start. Since the recognition of real activity location is out of the scope of this research, we simply check on the personal profile data that the respondents supported. In addition, we also use the prompt recall data to double-confirm the exact location and obtain the coordinates through a geo-coding process.
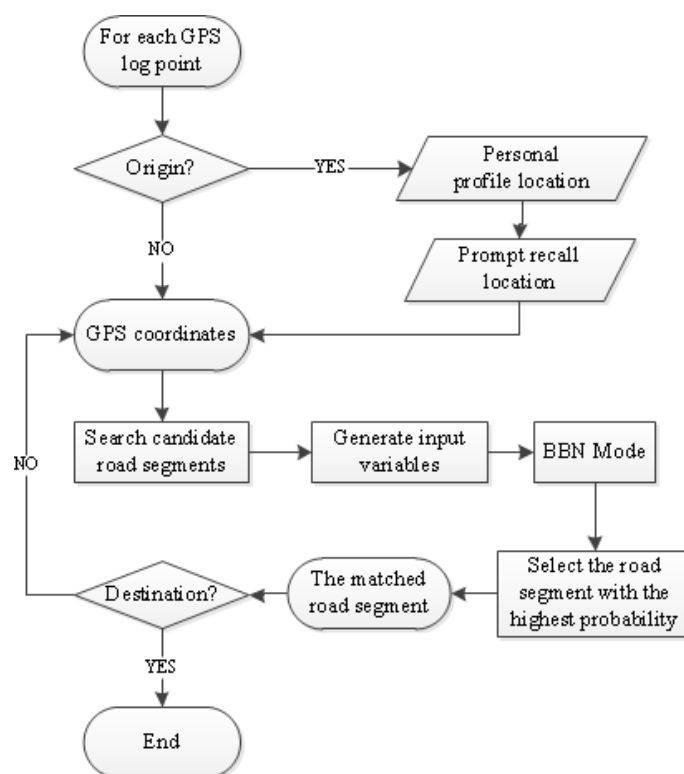

Figure 6 Flowchart of the implementation steps

After the recognition of the origin point, the candidate road segments are fetched and

relevant input variables are generated. The selection of the road segment follows the prediction of the BBN model. The selected road will be the one with a highest probability among all the candidate road segments. In this sense, the un-selected road segments also have a value of probability, which on the other hand measures the uncertainty of the map matching algorithm.

In addition, it should be noted that, for each GPS log points, the connectivity is considered here as one of the input variables to identify the road segment. This is different from the hard constraint which only considers connected road segments and excludes the potential roads which are not connected with the previous road segment. In some cases, it is possible that a road segment is incorrectly matched or a road segment is too short to be matched with any locations. Therefore, relaxing the connectivity as a hard constraint will avoid the possible error propagation in case that one road segment was recognized incorrectly. In order to ensure that the matched road segments are sections of a same route, we post process the matched data by checking from the origin to the destination. For example, if there is one road segment which does not connect with both the previous and the forward road segments, the end node of the previous road will be connected with the start node of the forward road by using the shortest path algorithm.

4. RESULTS

Because for each log point in the training dataset, a set of candidate road segments were gathered in advance, in which each road segment has a label on whether it is a matched road (true) or not (false). The inter-dependency correlations among input and output variables are then set up in terms of the ground truth. Then the data can be used for training the model to obtain the potential interdependency correlations which can be used for prediction.

Table 1 shows an example of the conditional probabilities for the connectivity variable specifically. The percentage of rightly matched is the relative value by comparing the number of data records which are matched correctly divided by the total sample. It represents the relative accuracy based on the current dataset. As one can see that the percentage of rightly matched road reaches 0.832, if the current road connects with the previous road, indicating that the connectivity variable is important especially in the case that the matched road belongs to the node.

For these roads that are in the candidate road set (here, the label is False) but not the matched roads, the connectivity results into a percentage of 0.465 and 0.535 for connected and unconnected road segments, respectively. This means the unconnected roads are more possible (0.535 vs. 0.465) to be un-matched road than matched roads, indicating the connectivity correlation will benefit to matching the road correctly.

Table 1 Conditional probability of connectivity

		Connect with the previous road (Connectivity)	
		Yes	no
The rightly matched road	True	0.832	0.168
	False	0.465	0.535

In order to evaluate the performance of the proposed BBN model, we use the results according to the indicators of the correctly classified instances (CCI), incorrectly classified instances (ICI) and Kappa value (Kappa). The ICI, being the other round of CCI, labels the percentage of incorrectly matched data. The kappa statistic measures the agreement of prediction with the true class, with 0 and 1 signifies incomplete and complete agreement, respectively. A higher value of Kappa indicates a better performance of the model.

Table 2 presents the details of the prediction accuracy. It is found that the accuracy of correctly classified instances 87.02%. This means the BBN model has a good performance in matching the GPS data. The level of Kappa for BBN model is 0.577, which is also satisfied in this context.

Table 2 Prediction accuracy and model performance

	CCI	ICI	Kappa
All samples	87.018%	12.983%	0.577

Table 3 shows the results of hit ratio for the two classes of the output variable in the BBN model. Hit ratio is a measure of business performance traditionally associated with sales. It is normally a matrix which includes all the combinations of different classes in the prediction result. Here, the hit ratio shows how accurate a road was matched. The higher is the value, the more accurate are the matched results. Since we set the variable into 2 levels, one is the right road segment, and the other is the found road segments but not the right road. We found for both classes, the accuracy percentages are 56.19% and 95.64%, respectively. These ratios are calculated based on the whole candidate road segments, which means that the more number of candidate road, the lower of the level of confirmed ratio, because only one road among the candidate roads can be labeled as the right road. This is why the ratio of confirmed yes is lower than that of the confirmed no, because the number of unmatched roads in the candidate road set is large.

Table 3 Results of the hit ratio

	Confirmed yes	Confirmed no
Hit ratio	56.19%	95.64%

5. SUMMARY, CONCLUSIONS AND DISCUSSION

Identifying the location of a GPS point on a road segment has been an important research issue in transportation planning, transport modeling and environmental analysis. In route choice simulations, the GPS data provides the merit with more accurate and detailed real-time information that the traditional state choice experiment cannot obtain. The matched roads in combination with the time information will also contribute by providing large opportunities in microscopic simulation of traffic flow and environmental analyses.

Development of an efficient map matching algorithm is of high importance in different research fields. Due to the fact previous methods are not efficient enough in the aspect of flexibility and learning capability, in this paper, we proposed an algorithm using Bayesian belief network model for map matching of GPS data. Variables employed in the model are distance to road, difference of directions, difference of angles of two adjacent links,

connectivity, number of satellites and PDOP. The algorithm is investigated by using the data collected in the Eindhoven region, The Netherlands. Simulation results showed that the BBN model shows a good performance in recognizing the road segments with an accuracy of 87% (correctly classified instances).

As one may argue that the accuracy of GPS data is questionable, which are a common issue in this study as well as others related to map matching algorithms, such data should not give main influences on the matching results. Although the extreme unrealistic data appear not much frequent in our GPS data, it is always true that the effect of the noise data should be handled well. Therefore, when applying this algorithm in real applications, a process to filter such data out is necessary. Moreover, the algorithm needs to be improved by designing it more flexible with respect to these inaccurate data.

As presented in the paper, the performance of the algorithm was evaluated through a carefully recorded dataset. This was considered as a feasible way to confirm the validity of this algorithm, because the validation of an algorithm in the context of map matching is not much straightforward. There needs additional effort to provide example data where sufficient ground truth and various special cases should be included. In spite of the acceptable efficiency of Bayesian network in map matching (87%), one cannot say it is a better algorithm than others at this moment because of the difference in the GPS data used. Future research will go further to examine some other algorithms by using the same data set to see the possible superiority of the proposed algorithm.

Since this paper is considered as a prototype of an enhanced map matching algorithm, here, we only included a small sample to test the efficiency of this algorithm. Future work will include more samples to further check the generality in large scale applications. In addition, some special cases which are common in GPS data, like passing through a cross-section, U-turn, driving through round-about and traveling off-roads, are necessary to give a special care. Some facts along with these special cases would then be necessary to extract in advance from the GPS data.

REFERENCES

Bernstein, D., Kornhauser, A. (1998) An introduction to map matching for personal navigation assistants. http://www.njtude.org/reports/mapmatchintro.pdf.

Du, J., Aultman-Hall, L. (2007) Increasing accuracy of trip rate information from passive multi-day GPS travel datasets: automatic trip end identification issues. Transportation Research Part A, 41, 2020-232.

Feng, T., Timmermans, H.J.P. (2012) Recognition of transportation mode using GPS and accelerometer data. Paper presented at the 12th International Conference on Travel Behavior Research, Toronto, Canada, July 2012.

Lin, L., Patterson, D.J., Fox, D., Kautz, H. (2007) Learning and inferring transportation routines. Artificial Intelligence, 171, 311-331.

Michel, B., Chen, J., Newman, J. (2013) A probabilistic map matching method for smartphone GPS data. Transportation Research Part C: Emerging Technologies 26, 78-98.

Mohammed A. Quddus, M.A., Noland, R.B., Ochieng, W.Y. (2006) A High Accuracy Fuzzy Logic Based Map Matching Algorithm for Road Transport. Journal of Intelligent Transportation Systems, 10(3), 103-115.

Moiseeva, A., Jessurun, J. and Timmermans, H.J.P. (2010) Semiautomatic Imputation of Activity Travel Diaries. Transportation Research Record: Journal of the Transportation

Research Board, 2183, 60-68.

Pyo, J.S., Dong-Ho, S., Tae-Kyung, S. (2001) Development of a Map Matching Method Using the Multiple Hypothesis Technique. Intelligent Transportation Systems, Proceedings. IEEE, V 23-27.

Quddus, M.A., Noland, R.B., Ochieng, W.Y., (2006) The effects of navigation sensors and digital map quality on the performance of mapmatching algorithms. Presented at the Transportation Research Board (TRB) Annual Meeting of the Transportation Research Board, Washington D.C., January 2006.

Quddus, M.A., Ochieng, W.Y., Noland, R.B. (2007) Current map-matching algorithms for transport applications: State-of-the art and future research directions. Transportation Research Part C: Emerging Technologies 15(5), 312-28.

Ren, M (2012) Advanced map matching technologies and techniques for pedestrian/wheelchair navigation. Ph. D dissertation, School of Information Sciences, University of Pittsburgh.

Rudloff, C. and Ray, M. (2010) Detecting travel modes and profiling commuter habits solely based on GPS data. Transportation Research Board, January 10, Washington D.C.

Schuessler, N. and Axhausen, K.W. (2009) Map-matching of GPS traces on high-resolution navigation networks using the Multiple Hypothesis Technique. Technique (MHT), Working paper, 568, Institute for Transport Planning and System (IVT), ETH Zurich, Zurich.

Schuessler, N. and K.W. Axhausen (2009) Map-matching of GPS traces on high-resolution navigation networks using the Multiple Hypothesis Technique (MHT), Working paper, 568, Institute for transport planning and system (IVT), ETH Zurich, Zurich.

White, C.E., Bernstein, D. and Kornhauser, A.L. (2000) Some map matching algorithms for personal navigation assistants. Transportation Research Part C: Emerging Technologies 8, 91-108.

Wolf, J., Hallmark, S., Oliveira, M., Guensler, R., Sarasua, W. (1999) Accuracy issues with route choice data collection by using global positioning system. Transportation Research Record, 1660, 66-74.