Searching for Similar Trajectories on Road Networks using Spatio-Temporal Similarity

Jung-Rae Hwang[1], Hye-Young Kang[2], and Ki-Joune Li[2]

[1] Department of Geographic Information Systems, Pusan National University, Korea
[2] Department of Computer Science, Pusan National University, Korea
{jrhwang,hykang}@isel.cs.pusan.ac.kr, lik@pnu.edu

Abstract. In order to search similar moving object trajectories, the previously used methods focused on Euclidean distance and considered only spatial similarity. Euclidean distance is not appropriate for road network space, where the distance is limited to the space adjacent to the roads. In this paper, we consider the properties of moving objects in road network space and define temporal similarity as well as spatio-temporal similarity between trajectories based on POI (Points of Interest) and TOI (Times of Interest) on road networks. Based on these definitions, we propose methods for searching for similar trajectories in road network space. Experimental results show the accuracy of our methods and the average search time in query processing.

Keywords: Trajectories, Road Network Space, Spatio-Temporal Similarity

1 Introduction

In the real world, most moving objects exist in road network space rather than in Euclidean space. Nevertheless, most previous studies on moving object trajectories have been based on Euclidean distance. Euclidean distance is not appropriate for road network space in query processing or measuring similarity between moving object trajectories.

The previously used methods related to searching for similar moving object trajectories have several problems. First, the previous methods were based on Euclidean space. This is not suitable for road network space with the distance defined along a road. That is, it is difficult to apply the distance of Euclidean space to road network space. Second, the previous methods considered only spatial similarity without considering temporal similarity to search for similar moving object trajectories. For example, if two moving objects pass through the same points at different time intervals, we can know that they are similar to each other spatially, but not spatiotemporally.

In order to solve the problems of the previous methods, we investigate the properties of moving objects on road networks and consider spatial similarity as well as temporal and spatio-temporal similarity. In terms of real applications, we are not interested in meaningless locations or times. We consider interesting points on road networks. We also interest in the time interval of moving objects on road networks. In this paper, we define temporal similarity as well as spatio-temporal similarity between moving object trajectories based on POI (Points of Interest) and TOI(Times of Interest). Based on these definitions, we propose methods for searching for similar trajectories of moving objects on road networks.

This paper is organized as follows. In section 2, we introduce the related work and the problems of the previously used methods and propose the motivations of this paper. In section 3, we propose methods in order to search for similar trajectories on road networks based on POI and TOI. Experimental results are given in section 4. Finally, we conclude and suggest future work in section 5.

2 Related Work and Motivation

In this section, we introduce related work with moving object trajectories on road networks. We discuss the problem of previous methods and propose the motivations of this paper.

2.1 Related Work

In order to analyze the behavior of moving objects, we must first define a similarity measure between moving object trajectories. To define this similarity, research representing the trajectory of moving objects in road network space is required. There has been some research representing moving object trajectories. Models for representing and querying moving objects on road networks were presented in [1][2] and approaches for representing and reasoning moving objects in constrained environments moving along a road network were introduced in [3][4].

In previous studies, the similarity measure between trajectories was based on Euclidean space and considered only spatial similarity without considering temporal and spatio-temporal similarity. In particular, some methods searching for similar trajectories were introduced in [5][6]. The method proposed in [6] searched for the most similar trajectory with a given query trajectory. However, it is not suitable for road networks because this method is based on Euclidean space. The similarity retrieval for the trajectory of mobile objects was presented in [7], which determined the similarity between trajectories based on shape. Contrary to other existing studies, this considered the spatio-temporal similarity but has the problem of Euclidean distance. Methods searching for similar trajectories using the distance function based on OWD (one way distance) or Time Warping Distance were proposed in [8][9]. These methods also considered only spatial similarity without considering temporal and spatio-temporal similarity and it is difficult to apply these to road network space because of the problem of Euclidean distance and Time Warping Distance.

2.2 Motivation

In order to search moving object trajectories, some methods of existing research have been proposed in Euclidean space. However, Euclidean distance is not appropriate for road network space defined along a road. We investigate several differences between Euclidean space and road network space. First, while moving object trajectories in Euclidean space are expressed to a sequence of points in (x,y,t) space, those of road networks are represented as a set of (*SegID*,*offset*,t), where *SegID* is a road sector identifier, and *offset* is the offset from the starting point of the road sector. Therefore, the distance between two points is calculated more easily on road networks defined along road sectors than Euclidean space. Second, moving object trajectories in Euclidean space have a linear interpolation problem.

Figure 1 shows the difference in linear interpolation between Euclidean space and road network space. In figure 1, the moving object trajectory TR_A passes a and c. For example, suppose that find an intermediate point between a and c. Then, we can find point b in road network space using (*SegID*, *offset*,t), but find point b' in Euclidean space. This means that Euclidean distance is not suitable for road network space.

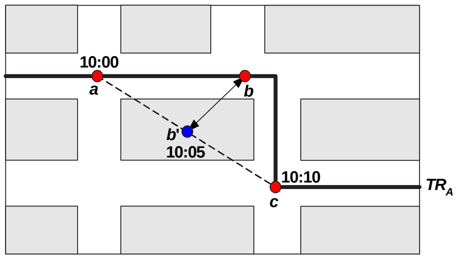

Fig. 1. Linear interpolation between Euclidean space and road network space

Based on these properties, figure 2 shows another example of the difference between Euclidean space and road network space. In figure 2, suppose that find the nearest two gas stations from a moving vehicle. If we find them by Euclidean distance, they are a and d. However, if we find them along a road, they are e and f. Consequently, we can know that a and d are meaningless gas stations from a given vehicle.

Most previous methods considered only spatial similarity in measuring the similarity between moving object trajectories. For example, if two trajectories pass through the same points at different time intervals on road networks, we understand by spatio-temporal intuition that they are not similar to each other. However, previous methods asserted that two trajectories are similar to each other.

To solve these problems concerning previous methods, we define spatial and temporal similarity based on road networks. In general, moving objects on road

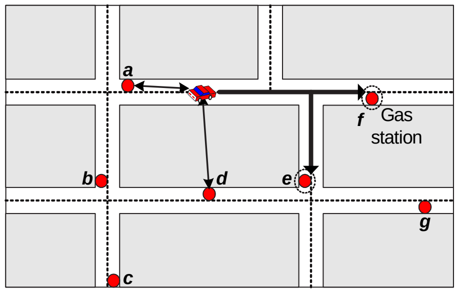

Fig. 2. Difference between Euclidean distance and road network distance

networks are represented as locations and times obtained by GPS. In real applications, however, we are not interested in these meaningless locations or times. We consider POI (Points Of Interest) on road networks and TOI (Times Of interest) of moving objects. In this paper, we define similarity between moving object trajectories based on POI and TOI and search for similar trajectories based on this similarity. For example, if two trajectories passed the same POI and TOI, we assert that they are similar to each other.

Figure 3 shows an example of how to define similarity of trajectories based on POI and TOI. For example, suppose that compare the similarity $Sim(TR_A,TR_B)$ between two trajectories TR_A and TR_B and the similarity $Sim(TR_B,TR_C)$ between two trajectories TR_B and TR_C. In figure 3, while TR_A and TR_B have a few temporal differences because they pass the same POIs at different time intervals, TR_B and TR_C do not fully pass the same POIs but pass at an almost similar time interval. In figure 3, it is hard to distinguish the differences between $Sim(TR_A,TR_B)$ and $Sim(TR_B,TR_C)$. Therefore, we define not only temporal similarity and spatio-temporal similarity but also spatial distance and spatio-temporal distance based on POI and TOI. We propose methods for searching for similar trajectories using these definitions.

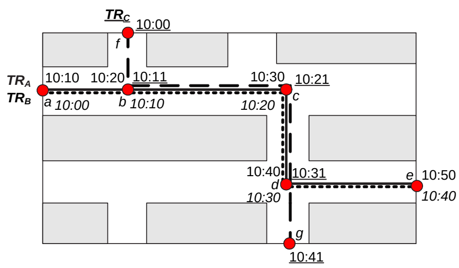

Fig. 3. How to define similarity of trajectories

3 Methods for Searching for Similar Trajectories on Road Networks

It is difficult to search directly for similar trajectories from a number of trajectories on road networks. To search for similar trajectories on road networks, therefore, we need a filtering step. In this paper, we use spatial filtering and temporal filtering because we search for similar trajectories based on spatial and temporal similarity. Spatial filtering is considered by spatial similarity, which is based on POI. The spatial similarity was proposed in our previous research[10]. Temporal filtering is considered by temporal similarity, which is based on TOI. For example, if two trajectories passed by the same TOI, it is regarded that they are similar to each other temporally.

After performing the filtering step based on the spatial or temporal similarity, we need a refinement step in order to search for similar trajectories. In this paper, we use the refinement step in order to search for similar trajectories from the trajectories selected by the filtering step. Figure 4 shows the process of searching for similar trajectories by using the filtering step and the refinement step.

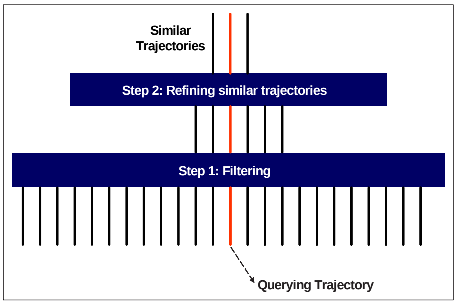

Fig. 4. Process of searching for similar trajectories

We introduced the following three methods in our previous research in order to search for similar trajectories on road networks[10]. Among three methods, we proposed the method for searching for similar trajectories in [10], which consists of two steps; the filtering step based on spatial similarity and the refinement step based on temporal distance. We also investigated problems related to the remaining methods. In this paper, we present methods in order to search for similar trajectories by solving the problems of our previous research.

- Method 1: Searching for similar trajectories based on spatial filtering and temporal distance.
- Method 2: Searching for similar trajectories based on temporal filtering and spatial distance.
- Method 3: Searching for similar trajectories based on spatio-temporal filtering and spatio-temporal distance.

In this section, we investigate each method in detail.

3.1 Searching for Similar Trajectories based on Spatial Filtering and Temporal Distance

We proposed method 1 in our previous research. We briefly introduce method 1 in this subsection. Method 1 filters trajectories based on spatial similarity and refines similar trajectories based on temporal distance. We defined the spatial similarity between trajectories based on POI in method 1.

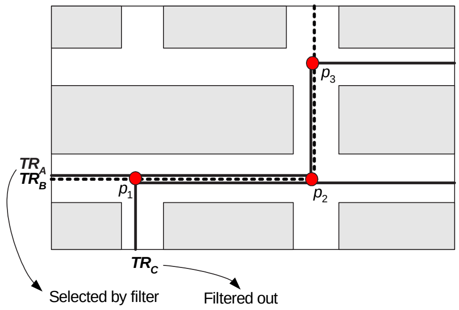

Fig. 5. Spatial Filter by POIs

Figure 5 shows an example of filtering based on the spatial similarity for POIs (p_1,p_2,p_3) of a given query. As shown in figure 5, TR_A and TR_B were selected by filtering and TR_C was filtered out.

In [10], we defined spatial similarity as well as the temporal distance between trajectories so as to apply method 1. Figure 6 shows an example of temporal distance between two trajectories. In this figure, the temporal distance between two trajectories TR_A and TR_B is calculated as follows:

$$dist_T(TR_A, TR_B, P) = 10 + 10 + 10 + 20 + 15 = 65$$

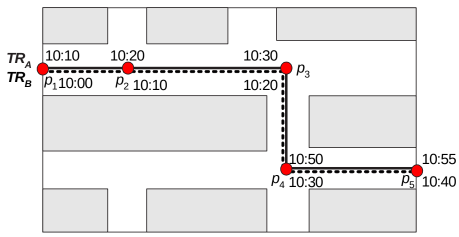

Fig. 6. An example of temporal distance

The advantage of this method is that a lot of trajectories are removed from trajectory data by spatial filtering. The disadvantage is that a long period of time is required in spatial filtering because the time complexity for the comparison between POIs of the query trajectory and those of trajectory data is $O(n^2)$.

3.2 Searching for Similar Trajectories based on Temporal Filtering and Spatial Distance

In terms of practical application, the meaning of distance between two time intervals can rarely be found. Thus, our previous research introduced that method 2 was not appropriate for searching for similar trajectories. However, we are interested in time intervals of moving objects. TOI (Times of interest) is an important characteristic of road networks. If trajectories pass the same points at the same TOI on road networks, we consider that they are similar to each other. Therefore, we define temporal similarity based on TOI. For example, the heaviest traffic time intervals on a specific road network can be TOI. We filter trajectories using this definition. If two trajectories pass through the same TOI, they are considered similarity by the following definition:

Definition 1. *Temporal Similarity between Trajectories on Road Networks Suppose that T is a set of TOIs on a given road networks. Then, temporal similarity between two trajectories TR_A and TR_B is defined as*

$$Sim_{TOI}(TR_A, TR_B, T) = \begin{cases} 1, & if \forall t \in T, \\ & \quad t \in [t_s(TR_A), t_e(TR_A)] \ \&\& \ t \in [t_s(TR_B), t_e(TR_B)] \\ 0, & otherwise \end{cases}$$

Figure 7 shows an example of filtering based on temporal similarity. Suppose that TOI of a given query trajectory is [8:00 $\sim$ 9:00]. Then, TR_C, TR_D and TR_E are selected according to temporal similarity.

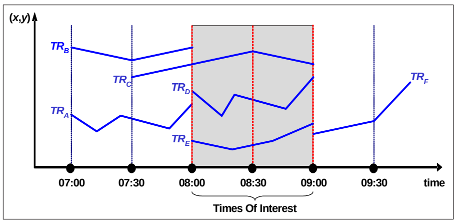

Fig. 7. Temporal Filter by TOI

With filtered trajectories based on temporal similarity, we define spatial distance and refine similar trajectories based on this. Spatial distance can be defined as the difference between the locations of two objects passing the same TOI as follows:

Definition 2. *Spatial Distance between Trajectories*
Suppose that $t \in T$, and T is the set of TOIs. Then the spatial distance between two trajectories TR_A and TR_B is defined as

$$dist_S(TR_A, TR_B, T) = \sum dist_S(p(TR_A, t_i), p(TR_B, t_i))$$

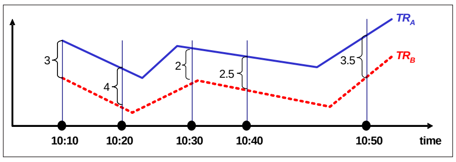

Fig. 8. An example of spatial distance

Figure 8 shows an example of the spatial distance between two trajectories TR_A and TR_B. Their spatial distance is calculated as follows:

$$dist_S(TR_A, TR_B, T) = 3 + 4 + 2 + 2.5 + 3.5 = 15$$

Algorithm 1. Searching on based Temporal Filter and Spatial Distance

---

Input. input trajectories TR_{IN}, threshold δ, query trajectory tr_Q, TOI set T, time interval t
Output. similar trajectories TR_{OUT}
Begin
  $TR_{Candidate} \leftarrow \phi$
  $TR_{OUT} \leftarrow \phi$
  For each $tr \in TR_{IN}$
    If $tr.t \supseteq tr_Q.t$
    then $TR_{Candidate} \leftarrow TR_{Candidate} \cup \{tr\}$
  For each $tr \in TR_{Candidate}$
    If $dist_S(tr_Q, tr, T) < \delta$
    then $TR_{OUT} \leftarrow TR_{OUT} \cup \{tr\}$
  return TR_{OUT}
End

Consequently, method 2 searches for similar trajectories using filtering based on temporal similarity and refining based on spatial distance. However, the disadvantage of this method is that many trajectories are selected from trajectory data by temporal filtering. For example, if the time interval of a query trajectory is much shorter than the total time interval for all moving objects, most trajectories are selected from trajectory data; nevertheless, the advantage is that little time is required in temporal filtering because the comparison between the time interval of a query trajectory and that of the trajectory data can be calculated simply and quickly.

Algorithm 1 summarizes the search procedure of method 2 explained in this subsection. It consists of two steps; the filtering step based on temporal similarity and the refinement step used in order to search for similar trajectories based on spatial distance.

3.3 Searching for Similar Trajectories based on Spatio-Temporal Filtering and Spatio-Temporal Distance

There is a possibility that we search more efficiently for similar trajectories, if spatio-temporal similarity is considered in searching for similar trajectories on road networks. Therefore, method 3 considers both methods 1 and 2. That is, method 3 uses spatial and temporal similarity together in the filtering step. Afterwards, we refine similar trajectories using spatio-temporal distance based on POI and TOI. In order to apply this method, we need a definition for measuring spatio-temporal distance. However, we stated in section 2 that it is difficult to define similarity between trajectories by spatio-temporal distance directly. In this paper, we regard spatio-temporal distance as the sum of temporal distance and spatial distance, which is defined as follows:

Definition 3. *Spatio-Temporal Distance between Trajectories*
Suppose that TR_A and TR_B are two trajectories. Then the spatio-temporal distance between TR_A and TR_B is

$$dist_{ST}(TR_A, TR_B) = dist_T(TR_A, TR_B) + dist_S(TR_A, TR_B)$$

To use this definition, the equivalence between temporal distance and spatial distance is defined so that 1 second = α meters. Moving objects on road networks move with various speeds. With this observation, we solve the equivalence problem between temporal distance and spatial distance using the speed of moving objects. That is, the equivalence problem between temporal distance and spatial distance is solved by the following formula:

$$Convt_S(TR_A, TR_B) = |(V_{TR_A}) - V_{TR_B})| \times dist_T(TR_A, TR_B)$$

The above formula converts temporal distance into spatial distance. Applying this formula to definition 3, the spatio-temporal distance between two trajectories is defined as follows:

$$dist_{ST}(TR_A, TR_B) = Convt_S(TR_A, TR_B) + dist_S(TR_A, TR_B)$$

By solving the equivalence problem, it is possible to represent the spatio-temporal distance as the spatial distance. Consequently, we search for similar trajectories based on spatiotemporal similarity and spatiotemporal distance. Figure 9 shows an example of the spatiotemporal distance between a query trajectory TR_{Query} and the other trajectory TR_A. In this figure, suppose that the distance from p_1 to p_4 is 30 km and that between each POIs is 10 km, with each speed of TR_{Query} and TR_A during the blocks being 60 km/h and 30 km/h. Then, a query trajectory passes through four POIs(p_1,p_2,p_3,p_4) during the time interval [10:00∼10:30].

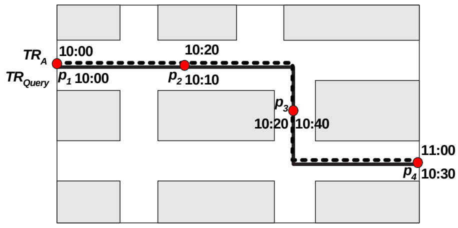

Fig. 9. An example of spatio-temporal distance

As shown in figure 9, TR_A satisfied the condition of TR_{Query}. Thus, the spatio-temporal distance between TR_{Query} and TR_A is calculated as follows by the above formula and definition:

$$\begin{aligned} dist_{ST}(TR_{Query}, TR_A) &= dist_T(TR_{Query}, TR_A) + dist_S(TR_{Query}, TR_A) \\ &= |(V_{Query} - V_A)| \times dist_T(TR_{Query}, TR_A) + dist_S(TR_{Query}, TR_A) \\ &= |(60km/h - 30km/h)| \times \text{ 60 minutes + 30 km = 60 km} \end{aligned}$$

The advantage of this method is that the similar trajectories with a query trajectory are selected by spatial and temporal filtering, but the disadvantage is that more similar trajectories than the selected trajectories are included among the trajectories removed by the filtering step. For example, figure 10 shows another example of the spatio-temporal distance between the query trajectory TR_{Query} and the trajectory TR_B removed by the filtering step. In this example, we follow the assumptions of figure 9. We just suppose that the speeds of two trajectories TR_{Query} and TR_B are 60 km/h and 75 km/h.

When apply equally with the example of figure 9, the spatio-temporal distance between two trajectories is calculated as follows:

$$\begin{aligned} dist_{ST}(TR_{Query}, TR_B) &= dist_T(TR_{Query}, TR_B) + dist_S(TR_{Query}, TR_B) \\ &= |(V_{Query} - V_B)| \times dist_T(TR_{Query}, TR_B) + dist_S(TR_{Query}, TR_B) \end{aligned}$$

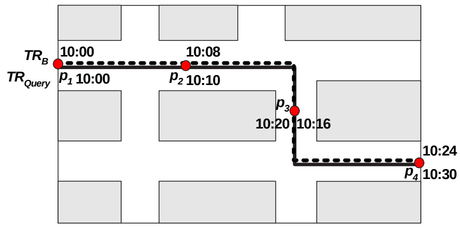

Fig. 10. Another example of spatio-temporal distance

$$= |(60km/h - 75km/h)| \times 12 \text{ minutes} + 15 \text{ km} = 18 \text{ km}$$

With $dist_{ST}(TR_{Query}, TR_A)$ and $dist_{ST}(TR_{Query}, TR_B)$ of the above examples, we know that TR_B is more similar than TR_A to the query trajectory TR_{Query}. However, method 3 does not compare TR_B with TR_{Query} because TR_B is a trajectory removed by the filtering step based on temporal similarity.

Algorithm 2 summarizes the search procedure of method 3 explained in this subsection.

Algorithm 2. Searching on based Spatio-Temporal Filter and Spatio-Temporal Distance

---

Input. input trajectories TR_{IN}, threshold δ and ε, query trajectory tr_Q, POI set P, TOI set T, time interval t
Output. similar trajectories TR_{OUT}
Begin
  $TR_{Candidate} \leftarrow \phi$
  $TR_{OUT} \leftarrow \phi$
  For each $tr \in TR_{IN}$
    If $(\forall p \in P$, p is on $tr)$ && $(tr.t \supseteq tr_Q.t)$
    then $TR_{Candidate} \leftarrow TR_{Candidate} \cup \{tr\}$
  For each $tr \in TR_{Candidate}$
    If $(dist_T(tr_Q, tr, P) < \delta)$ && $(dist_S(tr_Q, tr, T) < \varepsilon)$
    then $TR_{OUT} \leftarrow TR_{OUT} \cup \{tr\}$
  return TR_{OUT}
End

---

4 Experimental Results

In order to examine the feasibility of methods proposed in this paper, we performed experiments that compare the accuracy and the performance of our methods. In previous research, the most representative moving object generator based on road networks was T.Brinkhoff's moving object data generator[11][12]. Data generated by T.Brinkhoff's generator is not fit to real data because its movement is uniform and acceleration and deceleration are unexpressed. Thus, We experimented with a moving object generator based on real road networks in Pusan. This generator reflects the real road information and the traffic information and generates a near real moving object data by adding the various speeds of moving objects.

Figure 11 shows the generator of moving object data used in this paper. We defined 10,000 POIs on road networks in Pusan. We generated 100,000 moving object trajectories using this generator and 5,000 query trajectories from the moving object trajectories.

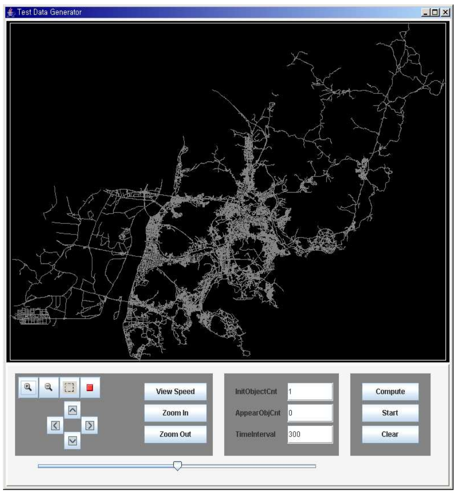

Fig. 11. Generator of moving object trajectories based on road networks

In our experiments, we compared our methods proposed in this paper. Figure 12 shows the consistency rate when searching for the same trajectory between method 1 and method 3. As shown in this figure, method 1 and method 3 show a high consistency rate. Here, we regarded the search result of the same trajectory by using these two methods as having been in agreement. For example, suppose that find the most similar trajectory with a query trajectory. As shown in this figure, the consistency rate between method 1 and method 3 is 100 % because they found the same trajectory.

We excluded method 2 from this experiment because method 2 searched for so many trajectories and they included meaningless trajectories. These meaningless trajectories are trajectories that pass different POIs with a query trajectory. This means that most trajectories are selected by temporal filtering because the time interval of a query trajectory is smaller than the total life span for all moving objects.

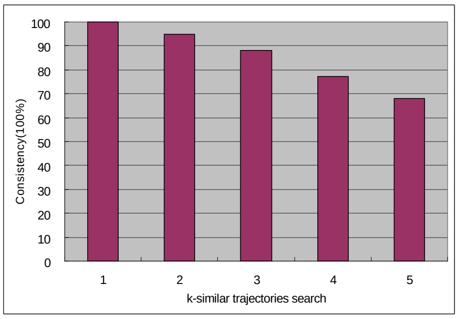

Fig. 12. Consistency rate between method 1 and method 3

Figure 13 shows the average search time in query processing for our methods by changing the number of POIs. As stated above, method 2 searched for so many trajectories including meaningless trajectories. As shown in this figure, however, method 2 required the least search time. Contrary to our expectations, this means that method 1 is more necessary the search time than method 2. Because the time complexity of the spatial distance and the temporal distance is $O(N)$, those of the spatial filtering and the temporal filtering are $O(n^2)$ and a constant.

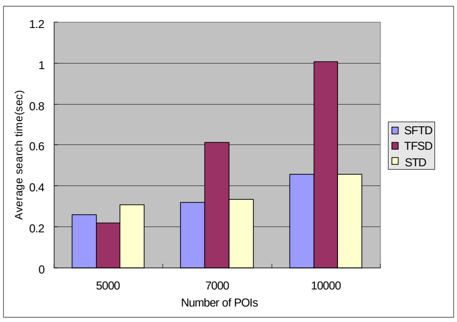

Fig. 13. Average search time in query processing

5 Conclusion and Future Work

The previously used methods related to searching for similar moving object trajectories were based on Euclidean space. A few studies introduced a similarity measure between moving object trajectories in road network space, but most studies considered only the spatial similarity between trajectories. In this paper, we defined the temporal similarity and the spatio-temporal similarity as well as the spatial distance and the spatio-temporal distance based on POI and TOI. Based on these definitions, we proposed methods for searching for similar trajectories on road networks. Our experimental results showed the accuracy of our methods and the average search time in query processing.

In the future, we will apply data mining techniques such as pattern analysis or clustering method to moving object trajectories on road networks. By using these mining techniques, detection of specific patterns of moving object trajectories or clusters of similar trajectories may be possible in the future.

Acknowledgements

This research was supported by the Internet information Retrieval Research Center(IRC) in Hankuk Aviation University. IRC is a Regional Research Center of Kyounggi Province, designated by ITEP and Ministry of Commerce, Industry and Energy.

References

1. Michalis Vazirgiannis and Ouri Wolfson. A Spatiotemporal Model and Language for Moving Objects on Road Networks. In *Proceedings of the Seventh International Symposium on Spatial and Temporal Databases*, pages 20–35. Springer-Verlag, 2001.
2. Laurynas Speicys, Christian S. Jensen, and Augustas Kligys. Computational Data Modeling for Network-constrained Moving Objects. In *Proceedings of the Eleventh ACM International Symposium on Advances in Geographic Information Systems*, pages 118–125, 2003.
3. Nico Van de Weghe, Anthony G. Cohn, Peter Bogaert, and Philippe De Maeyer. Representation of Moving Objects along a Road Network. In *Proceedings of the twelfth International Conference on Geoinformatics*.
4. Nirvana. Meratnia and Rolf A. de By. Representation in Location-based Services: Problems & Solution. In *Proceedings of the Third International Workshop on Web and Wireless Geographical Information Systems*.
5. Michail Vlachos, Dimitrios Gunopulos, and George Kollios. Robust Similarity Measures for Mobile Object Trajectories. In *Proceedings of the Thirteenth International Workshop on Database and Expert Systems Applications*, pages 721–728. IEEE Computer Society, 2002.
6. Michail Vlachos, George Kollios, and Dimitrios Gunopulos. Discovering Similar Multidimensional Trajectories. In *Proceedings of the Eighteenth International Conference on Data Engineering*, pages 673–684. IEEE Computer Society, 2002.

7. Yutaka Yanagisawa, Jun ichi Akahani, and Tetsuji Satoh. Shape-Based Similarity Query for Trajectory of Mobile Objects. In *Proceedings of the Fourth International Conference on Mobile Data Management*, pages 63–77. Springer-Verlag, 2003.
8. Bin Lin and Jianwen Su. Shapes based Trajectory Queries for Moving Objects. In *Proceedings of the Thirteenth ACM International Workshop on Geographic Information Systems*, pages 21–30, 2005.
9. Choon-Bo Shim and Jae-Woo Chang. Similar Sub-Trajectory Retrieval for Moving Objects in Spatio-temporal Databases. In *Proceedings of the Seventh East-European Conference on Advances in Databases and Informations Systems*, pages 308–322. Springer-Verlag, 2003.
10. Jung-Rae Hwang, Hye-Young Kang, and Ki-Joune Li. Spatio-temporal Similarity Analysis between Trajectories on Road Networks. In *Perspectives in Conceptual Modeling, ER Workshop on CoMoGIS*, pages 280–289. Springer, 2005.
11. Thomas Brinkhoff. Generating Network-Based Moving Objects. In *Proceedings of the twelfth International Conference on Scientific and Statistical Database Management*, pages 253–255, 2000.
12. Thomas Brinkhoff. A Framework for Generation Network-Based Moving Objects. *GeoInformatica*, 6(2).