



Analysis of GPS Trajectories Mapping on Shape Files Using Spatial Computing Approaches

Saravjeet Singh^(✉) and Jaiteg Singh

Chitkara University Institute of Engineering and Technology, Chitkara University,
Rajpura 140401, Punjab, India
saravjeet.singh@chitkara.edu.in

Abstract. Locating the position of the device on the road network is a crucial component of a location-based system. The performance of location-based systems is highly affected by the mapping of user location on the digital map. Many spatial computing methods were developed by the research community to map the GPS fix on to the digital map. These mapping methods take GPS data and spatial data as input. The data used by these spatial computing algorithms is in huge amounts and is recorded using sensors and GPS receivers. While handling GPS data, these algorithms face many issues and based upon that each method has its advantages and disadvantages. This paper provides working of a few prominent methods. As per the research trends, four methods are considered and empirical evaluation and analysis are presented. To analyze the performance of these methods a standard data-set is used. 3 different routes having nodes in the range of 100 to 12598 are considered for this experiment. In this experiment mapping, results are analyzed using GPS trajectories of commutative distance 154.2 km. Performance and accuracy of considered map matching algorithms were analyzed on a total of 16804 GPS points. Results are analyzed using RMSE, accuracy ratio, and execution time. HMM-based map matching algorithm is considered as the most preferred algorithm having 94% accuracy with average execution time.

Keywords: GPS · Hidden Markov Model · Map matching · OpenStreetMap · Frechet distance

1 Introduction

As technology is advancing and smartphone become essential for many day to day activities. People are using mobile for banking, networking, entertainment, shopping, navigation, and many more. Navigation is one of the most using activities by the users using the smartphone. As per our observations, people primarily use smartphones for navigation. So there exists a huge demand for a mobile-based navigation system for route finding and route guidance activities. Spatial computing is a core component of the navigation system and is used to process spatial

data to find meaningful results. One simple example of spatial computing is map matching algorithms and is used to match the geographical information of the entity on a digital map and known as map matching algorithm. For the outdoor navigation system, map matching algorithms provide the mapping between data received from the GPS receiver with the spatial database. Main components of spatial computing are shown in Fig. 1.

Spatial Database. A spatial database is a database that provides spatial data types and spatial queries for the data storage and processing. It provides spatial indexing and spatial joins on the spatial data.

GPS Information. GPS information mainly deals with the temporal and positional information of the geographical entity. This positional information is used to find the actual location of the entity.

Spatial Computing Method. For navigation systems, map matching algorithms are main spatial computing methods. Map matching algorithm takes data from a spatial database, GPS receiver and optionally from (inertial sensors and other resources) and provides the actual location of the geographical entity.

User Output. This is actual output component which shows end result of the navigation system. It provides visual and audio output.

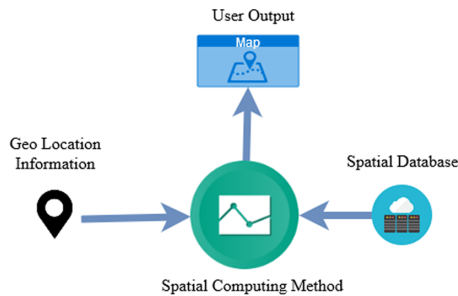


Fig. 1. Associated components in map matching process

1.1 Spatial Computing Techniques

Spatial computing techniques are main component of the navigation system. These acts as the heart of the navigation system. For this mapping application, we used map matching algorithms as a spatial computing technique. In a navigation system, these algorithms provide the actual location of the moving entity on the road network. As an output, these algorithms generate output on the GUI device. The main source of input for the map matching algorithm is shown in Fig. 2. Map matching algorithms are used for finding exact location the person

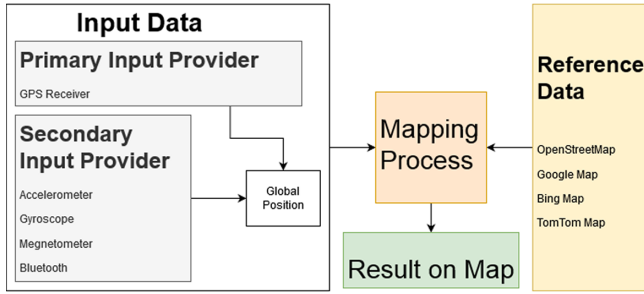


Fig. 2. Source of input for map matching algorithms

on road network and due to this, these algorithms are used in applications like route finding, geo-fencing, geo-tracking, distance between two location, shortest path problems, etc. Map matching algorithm uses two types of inputs; reference data and input data and are explained as below:

Input Data: Input data for map matching algorithm is a global position of a geographic entity (normally known as trajectories) and is fetched from GPS receiver, Bluetooth or inertial sensors. Below are the different type of input data providers:

- GPS receiver
- Accelerometer
- Gyroscope
- Magnetometer
- Bluetooth

Reference Data: Reference data is the actual data of the digital map. In map matching algorithm input data is mapped to reference data (shape file). Following are the type of reference data:

- OpenStreetMap
- Google Map
- Bing Map
- TomTom Map

Figure 3 demonstrate a typical map matching problem. A road network has n nodes. A vehicle is moving on the road, GPS receiver output (trajectory) is shown by red symbols. Location F is not on road network but it is away from the actual location. So map matching algorithm maps point F to the near by point on the road network. As in Fig. 3, point R1 and R2 are near to F and are on road network. So map matching algorithm provides best mapping of F with either R1 and R2.

Following are the few scenarios which cause poor performance of map matching algorithms [12, 13]:



Fig. 3. Map matching scenario

- Errors within positioning sensors
- complex urban road networks pose
- Absence of GNSS signal in Canyon regions
- Data quality issues in reference data
- Sampling rate of received GNSS data
- Odd values in input data

Many map matching algorithms have been proposed and used by the research community. Each algorithm used different approach, models and parameters to provide the better mapping result. Research is never-ending process and to proposed new algorithm we must aware of shortcoming in available tools and techniques. So in this paper, we attempted is to identify the shortcoming of prominent map matching algorithms and then using the shortcoming for the invention of new technique. In this paper we analyzed the mapping of GPS trajectories on shape files using different spatial computing techniques. The structure of this paper is as follows: Sect. 2 provides literature study of used spatial computing methods, detail of considered algorithm for evaluation are presented in Sect. 3, Sect. 4 provides detailed result analysis and discussion, Sect. 5 provide conclusion and future scope.

2 Literature Survey

Research for vehicular navigation was started in the early 1980's using the hardware device and simple small digitized map. After the popularity of GPS, research studies were done to utilize the GPS data to find the location of moving vehicles on digital map [4]. Many studies in the early 20's proposed many methods for spatial computing techniques. These methods used the concepts like road information, vehicle information, the direction of movement, dead reckoning, Kalman filter, filtration on GNSS signal, GNSS modeling, orthogonal projection, error correction, extended Kalman filter, etc. [1,2]. To enhance the accuracy of spatial computing using map matching process many approaches based on advanced model and filters came into existence. These approaches were

based on fuzzy logic, probability theory, path inference, sensor, etc. [3,7]. Probability theory played a very important role in the map matching process. Markov model (based on probability theory) provided a new direction to the map matching algorithm for online data [6,11]. Pattern and string matching techniques also used to enhance the performance of map matching algorithms [5]. Map matching using a genetic algorithm also provided good results [10]. To enhance the accuracy of the map matching algorithm, a dynamic two-dimensional algorithm was proposed [9]. Dijkstra's algorithm is further used to provide the map matching solution for the floating car data [8].

3 Considered Algorithm for the Spatial Computing

P2P: In the P2P method, a GNSS point is mapped to the closest node on the map. While implementing the P2P algorithm, only GNSS points and nodes of the road were considered. The perpendicular distance between the GNSS fix and nodes on the road were calculated. The point with minimum distance is calculated as a target matched point.

Topological Map Matching Algorithm: The topological map matching algorithm considers information of the road network, vehicle information, and GNSS fix to reconstruct the path. In this scenario, we considered road direction, road type, junction information, vehicle movement direction, and vehicle speed to fix the GNSS.

Hidden Markov Model Based Map Matching Algorithm: Hidden Markov model (HMM) based map matching algorithm uses the concept of the probability distribution to find the maximum likelihood between GPS fix and road network.

Frechet Distance Based Map Matching Algorithm: Frechet distance based map matching algorithm uses frechet distance between two-time variant curve to find the maximum likelihood between them. This algorithm uses free space diagram for finding the likelihood between two shape file and GPS trajectories.

4 Result Analysis

To analyze the performance and accuracy of map matching algorithms, an experiment is done. For the experimental study, we used P2P geometric, topological, HMM and frechet distance techniques. Three different routes were considered for the experiment and specifications of these routes are shown in Table 1. For test data (input data) collection, an android application was used to capture the GNSS receiver output after 1–5 s. 5 different users collected the data on different routes, their data collection details are shown in Table 2.

Total 5 GPS trajectories having 16804 GPS records were considered for the experiment. These trajectories were captured on the 3 routes. Reference data was selected from the OpenStreetMap dataset based on the selected routes. These

Table 1. Specification of considered route

Route no	Start address	End address	Distance
1	30.514851, 76.661137	30.300551, 76.844986	41 km
2	30.277279, 76.851552	29.946730, 76.821855	55.2 km
3	30.514925, 76.661524	30.518104, 76.658638	1.4 km

Table 2. Data collection detail

User no	Considered routes	Sampling interval	Collected GPS fix
1	1	5	3032
2	2	5	4198
3	2	2	9237
4	3	5	92
5	3	2	245

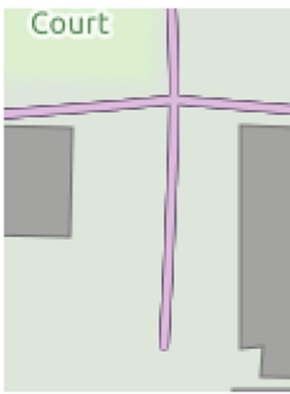
collected trajectories were mapped using the selected algorithms (output of these algorithms are shown in Fig. 4). Geometric and frechet based map matching algorithm provides incorrect results as shown in Fig. 4. Topological and HMM based map matching algorithm gives correct output as shown in Fig. 4(d) and (e).

In performed experiment map matching algorithm are analyzed by using matrices named Root Mean Square Error (RMSE), running time, and accuracy ratio. RMSE calculates the error of two datasets by comparing the mapping results as data series. Accuracy ratio calculates the mapping result by counting the correct and incorrect mapping points. According to performed experiment, HMM-based algorithm has less RMSE and high accuracy ratio whereas geometric algorithm has high RMSE and less accuracy ratio. Lower RMSE means high accuracy of the mapping. So according to RMSE value, HMM algorithm provides high accuracy. If we consider generalized result of analysis presented in Fig. 5 following is preference order of map matching algorithms:

1. HMM
2. Frechet based
3. Topological
4. Geometrical

Performance of an algorithm can be analyzed using execution time. Execution time is total time elapsed by an algorithm to provide the mapping result. Performance comparison of considered algorithms based upon the number of GPS points are shown in Fig. 6. According to performance comparison, P2P algorithm has low execution time. For less number of nodes, all algorithms have approximate same performance but at high node count, each algorithm has huge difference in execution time. If we consider generalized result of analysis presented in Fig. 6 following is preference order of map matching algorithm:

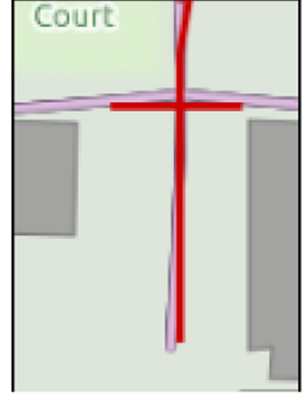
1. Geometrical
2. Topological
3. HMM
4. Frechet based



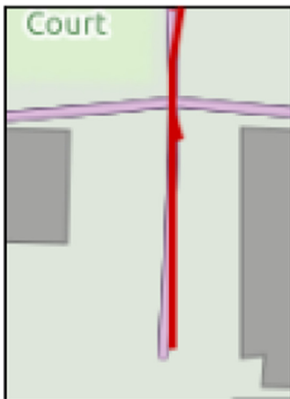
(a) Considered road network



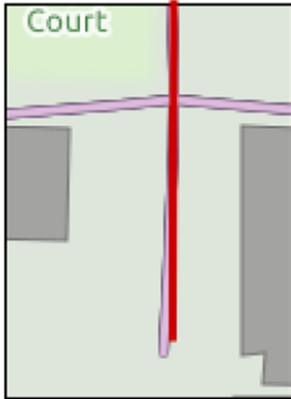
(b) Slow speed GPS data collected for map matching algorithm



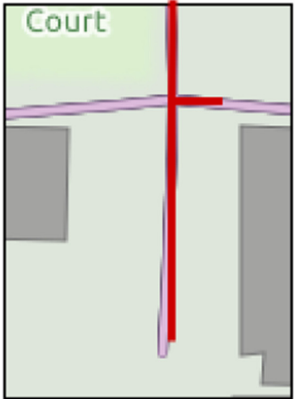
(c) P2P based map matching algorithm Output



(d) Topological map matching algorithm Output



(e) HMM based Map matching algorithm



(f) Frechet distance based Map matching algorithm

Fig. 4. Mapping output of four different selected map matching algorithms for same trajectory.

HMM based map matching algorithm have highest accuracy in comparison to other algorithm but average execution time. Topological algorithm has better

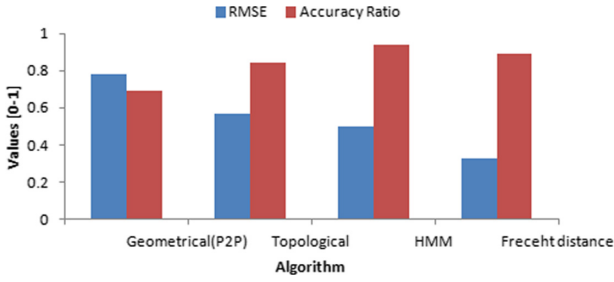


Fig. 5. Comparison of considered map matching algorithms based on RMSE and Accuracy ratio

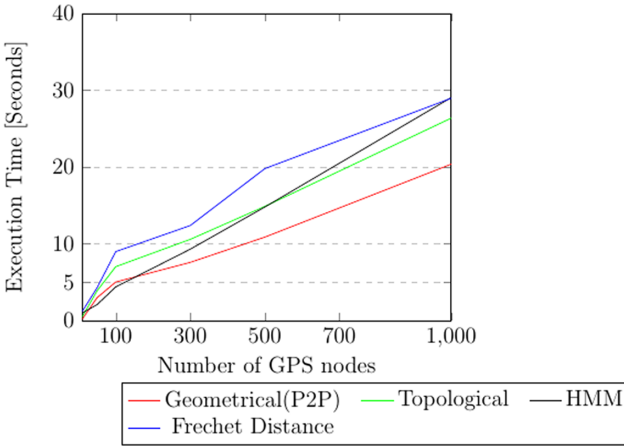


Fig. 6. Comparison of considered map matching algorithms based on execution time

accuracy in comparison to P2P algorithm, because topological algorithm considered topological information of both road network and vehicle. In topological algorithm the considered vehicle speed and direction of movement improves the algorithm accuracy at turn points and curved areas. Frenchet distance algorithm has high accuracy in comparison to P2P algorithm but require addition processing for free space calculation. Similarly P2P algorithm provides much faster result but has very low accuracy. Performance and accuracy both are important for map matching algorithm but we need to make a selection based on our requirement. Good accuracy and performance can not present in one algorithm so we need to select algorithm according to requirement. If we have sufficient processing power and space then we can select algorithm with higher accuracy. For scenario having low cost device with limited processing then preference should given to performance. According to combined analysis, if we considered average of both accuracy and performance then following is preference order to select an algorithm:

1. HMM
2. Topological
3. Frechet based
4. Geometrical

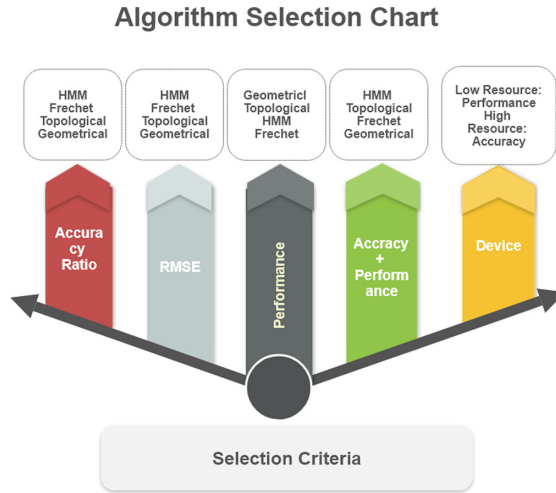


Fig. 7. Algorithm selection chart based on the user requirements and capabilities

5 Future Scope and Conclusion

In this paper, we analyzed the performance of spatial computing approaches (map matching algorithm) for mapping the GPS trajectories. Four prominent algorithms were selected and their empirical evaluation was presented. In this study, we considered point to point geometrical, topological, HMM-based, and frechet distance-based map matching algorithms. Performance and accuracy was evaluated and analyzed using the shape files of OSM data and GPS trajectories (collected using an android device). Test data was collected by five different users by traveling on three selected road network. GPS trajectories of five different users were mapped on OSM shape files. From the experimental results, it analyzed either algorithm has good performance or accuracy. For example, the HMM-based map matching algorithm has very high accuracy but poor performance and reverse is with P2P geometric map matching algorithm. As a conclusion, we presented a chart (as shown in Fig. 7) to select an algorithm based upon the user capabilities and requirement.

Accuracy and performance both are equally important but both at best range could not exist. So we need to make a selection. If the device has high processing power then accuracy should be preferred and if the device has low resources then, performance should be the preference. From this experiment, we concluded that

map matching algorithm being spatial computing method still requires improvement. Research is never-ending process, so this study will act as a base for the new algorithm that have enhanced performance and accuracy. In future we will implement the changes in any of the selected algorithms so as to enhance the performance and accuracy.

References

1. Bouju, A., Stockus, A., Bertrand, F., Boursier, P.: Location-based spatial data management in navigation systems. In: *Intelligent Vehicle Symposium 2002*, vol. 1, pp. 172–177. IEEE (2002)
2. Greenfeld, J.S.: Matching GPS observations to locations on a digital map. In: *Transportation Research Board 81st Annual Meeting* (2002)
3. Hunter, T., Abbeel, P., Bayen, A.M.: The path inference filter: model-based low-latency map matching of probe vehicle data. In: Frazzoli, E., Lozano-Perez, T., Roy, N., Rus, D. (eds.) *Algorithmic Foundations of Robotics X*. STAR, vol. 86, pp. 591–607. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-36279-8_36
4. Iwaki, F., Kakihara, M., Sasaki, M.: Recognition of vehicle's location for navigation. In: *Conference Record of Papers Presented at the First Vehicle Navigation and Information Systems Conference (VNIS 1989)*, pp. 131–138. IEEE (1989)
5. Miler, M., Todić, F., Ševrović, M.: Extracting accurate location information from a highly inaccurate traffic accident dataset: a methodology based on a string matching technique. *Transp. Res. Part C Emerg. Technol.* **68**, 185–193 (2016)
6. Mohamed, R., Aly, H., Youssef, M.: Accurate real-time map matching for challenging environments. *IEEE Trans. Intell. Transp. Syst.* **18**(4), 847–857 (2017)
7. Newson, P., Krumm, J.: Hidden Markov map matching through noise and sparseness. In: *Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pp. 336–343. ACM (2009)
8. Ptošek, V., Rapant, L., Martinovič, J.: Floating car data map-matching utilizing the Dijkstra's algorithm. In: Sharma, N., Chakrabarti, A., Balas, V.E. (eds.) *Data Management, Analytics and Innovation*. AISC, vol. 1016, pp. 115–130. Springer, Singapore (2020). https://doi.org/10.1007/978-981-13-9364-8_9
9. Sharath, M., Velaga, N.R., Quddus, M.A.: A dynamic two-dimensional (D2D) weight-based map-matching algorithm. *Transp. Res. Part C Emerg. Technol.* **98**, 409–432 (2019)
10. Singh, S., Singh, J., Sehra, S.S.: Genetic-inspired map matching algorithm for real-time GPS trajectories. *Arab. J. Sci. Eng.* **45**(4), 2587–2603 (2020)
11. Yang, C., Gidofalvi, G.: Fast map matching, an algorithm integrating hidden Markov model with precomputation. *Int. J. Geogr. Inf. Sci.* **32**(3), 547–570 (2018)
12. Zhao, F., Shin, J., Reich, J.: Information-driven dynamic sensor collaboration. *IEEE Signal Process. Mag.* **19**(2), 61–72 (2002)
13. Zhao, L., Ochieng, W.Y., Quddus, M.A., Noland, R.B.: An extended Kalman filter algorithm for integrating GPS and low cost dead reckoning system data for vehicle performance and emissions monitoring. *J. Navig.* **56**(2), 257–275 (2003)