Intelligent GPS trace management for human mobility pattern detection

Mahdi Hashemi

Cogent Engineering (2017), 4: 1390813
Intelligent GPS trace management for human mobility pattern detection

Mahdi Hashemi*

Abstract: Large volumes of volunteered GPS traces in the last decade have provided location-based services with an opportunity to become more intelligent and personalized. Individual and group mobility patterns, detected from GPS traces, can be used for this purpose. In this paper, we show the potential of GPS traces, if managed properly in the database, for detecting points of interest for individual users and even recognizing individual users from their walking patterns. However, when it comes to GPS traces, databases can be very complicated and cumbersome to populate. Databases provided by OSM and GeoLife do not effectively pave the path for data mining and machine learning techniques which require a much more detailed and organized database. A GPS trace database must provide statistics and detailed information about GPS traces not only for visualization purposes at the front-end, but also for cross checking purposes to eliminate erroneous records and to be applied in mobility pattern detection applications. This study provides the design of an interactive database management system for GPS traces whose applications in detecting points of interest and user identification are tested with GPS traces from the GeoLife project. The results show that while the accuracy of detected points of interest depends mostly on the size of data, the accuracy of user identification relies more upon the appropriate choice of input features to machine learning techniques.

Subjects: Machine Learning - Design; Database Management; Data Preparation & Mining; Intelligent & Automated Transport System Technology

Keywords: crowd-sourcing; pattern recognition; spatial-temporal data mining; intelligent transportation; artificial intelligence

ABOUT THE AUTHOR

Mahdi Hashemi received his PhD degree in Computing and Information from University of Pittsburgh in 2017. His research interests include theoretical machine learning, intelligent transportation systems, and computing with spatial-temporal data. He has published over 15 articles in top-tier journals.

PUBLIC INTEREST STATEMENT

People track themselves using their phones, as they walk, run, or driver. These spatial-temporal trajectories, collected in large numbers on a daily base, can be used to provide people with more intelligent internet services, such as wayfinding, shopping, and social community recommendations. This work explains how the appropriate management of these trajectories can lead to automatic knowledge discovery about people.
1. Introduction

Understanding the dynamics of large-scale human mobility patterns is beneficial to urban planning, traffic and transportation management, public transport design, emergency response management, public health, disease outbreak detection, and economic forecasting. The spread of mobile devices equipped with GPS receivers among people (Hashemi & Malek, 2012) has contributed to accumulation of large-scale GPS traces (Hashemi, 2017a; Hashemi & Karimi, 2014, 2016a, 2016b, 2017) which have motivated researchers from various fields to study human mobility (Liao, Patterson, Fox, & Kautz, 2007; Liu, Andris, & Ratti, 2010; Patterson, Liao, Fox, & Kautz, 2003; Zheng, Cao, et al., 2010). Song, Qu, Blumm, and Barabasi (2010) showed the high predictability in human mobility. Azevedo, Bezerra, Campos, and Moraes (2009) found that the movement velocity and acceleration of pedestrians follow a normal distribution. Lee, Hong, Kim, Rhee, and Chong (2009) effectively modeled human mobility using gaps among fractal waypoints. González, Hidalgo, and Barabási (2008) concluded that people tend to visit few locations frequently and highlighted the contrast between the simple repeated patterns in human mobility trajectories on one side and models such as Levy flight and random walk on the other side. Phithakkitnukoon, Lorenzo, Shibasaki, and Ratti (2010) found a strong correlation in daily activity patterns of people who share the same work area’s profile. Peng, Jin, Wong, Shi, and Lió (2012) developed a linear model to approximate the traffic flow between pairs of locations based on the experimentally inferred fact that people travel on workdays for three purposes: commuting between home and workplace, traveling from workplace to workplace, and others such as social activities. Li et al. (2012) used taxi traces to uncover patterns of pick-up quantity in urban hotspots and developed an ARIMA model to forecast how many passengers will be in a certain hotspot in the next time interval. Detecting points of interest (POIs) and identifying people through their walking patterns from large-scale GPS traces are two venues that we explore in our work.

However, GPS traces are stored in plain text formats with no attached metadata such as, transportation mode, length, or speed. This not only makes managing large volumes of GPS traces inefficient but also restricts the scale and scope of algorithms for human mobility pattern detection. Open Street Map (OSM), founded in UK in 2004 with more than 1 million registered users (Wood, 2013), is the most prominent volunteered geographic information devoted to providing a free map of the world emphasizing the road networks. Road networks are built upon GPS traces uploaded by registered users and can be edited or updated manually at any time. A description can be associated to a GPS trace while being uploaded but there are no additional required metadata or restrictions (https://www.openstreetmap.org/traces; OpenStreetMap, n.d.). This means the transportation mode of the GPS trace (e.g. walking, motoring, or boating) cannot be known in the database which in turn limits the database’s applications. Besides, they do not store additional metadata, such as average speed or total length of the GPS trace which can be automatically calculated. Such metadata not only facilitates analyzing, mining, and visualizing large volumes of GPS traces, but also paves the path for automatic applications of GPS traces. Examples of such applications are automatic road and pedestrian network construction (Hashemi, 2017b), recognizing POIs (Bhattacharya, Kulik, & Bailey, 2015), developing intelligent location-based services (Liu & Karimi, 2006), detecting individual (Song et al., 2010), or collective (Becker et al., 2013; Harder, Nes, Jensen, Reinau, & Weber, 2012) mobility patterns, and real-time event detection which is of great value to municipalities, police, and fire departments.

This paper shows how large volumes of GPS traces can be used to detect people’s mobility patterns, such as their POIs and to associate walking patterns with people’s identities. POIs can be used to make location-based or location-aware services more intelligent and personalized. For example, Patel, Chen, Smith, and Landay (2006) personalized the routes and shrank the navigation directions for drivers by applying their POIs. On the other hand, associating patterns in walking GPS traces with people’s identities can be used, for instance, in location-based social networks (e.g. friend recommendation (Yu, Pan, Tang, Li, & Han, 2011)) or in smart cities (Chang, Liu, Chou, Chen, & Shin, 2007; Ferrari, Rosi, Mamei, & Zambonelli, 2011; Pan et al., 2013) upon users’ permission. However, since detecting mobility patterns is not feasible without having sophisticated databases specifically designed for this purpose, we propose a structure for storing and managing crowd-sourced GPS traces.
2. GPS trace database management system

The entity relationship (ER) model and its corresponding relational logical model in Boyce-Codd normal form (along with functional dependencies) in the proposed GPS trace management system are represented in Figures 1 and 2. In Figure 1, rectangles show entities, rectangles with thick borders show weak entities, ovals show attributes, underlined attributes are primary keys, dashed underlined attributes are partial keys, arrows show key constraints (each entity appears in at most one instance of the relationship), and thick lines show total participation (all entities appear in at least one instance of the relationship). The attributes Owner_Username (representing the username of the person who uploaded this GPS trace) and Manager_Username (representing the username of the manager who is responsible for the user who uploaded this GPS trace) in GPS_TRACE table are foreign keys from tables USER and USER_MANAGER, respectively.

According to the ER, users can register in the system and upload GPS traces. Users are identified by their account username. A user can upload none or as many GPS traces as he/she wants. The user specifies the transportation mode for a GPS trace at the time of upload. Other information stored for a GPS trace include start time, total traveled distance, total time duration of the GPS trace, maximum and minimum longitude, maximum and minimum latitude, average longitude, average latitude, average altitude (some GPS traces may not contain altitude), and average speed. This information is not explicitly expressed in GPS traces and must be calculated. Each GPS trace has a set of GPS points. Each GPS point has a longitude, latitude, altitude, heading, HDOP (horizontal delusion of precision), date, time, and Unix time. Some points may not have altitude, heading, and HDOP. There are two types of managers: (a) user managers who manage a group of users, and (b) trace managers who manage a group of GPS traces.

![Figure 1. ER model of the GPS trace management system.](image1)

![Figure 2. Logical model of the GPS trace management system.](image2)
For the physical design, MySQL is used as the DBMS and Tomcat as the web server. The server is written in Servlets and the client in JSP. The post method is used for communication between the client and server. D3js is used to create charts in JSP pages. To produce those charts, JSON arrays are
constructed out of database query results in the server and sent to the client. The structure of our client-server system is shown in Figure 3. Arrows show the correspondence between clients and servers.

Figure 4 shows the JSP pages in a browser. After logging into the system, the user can see the user page which displays a list of the user’s GPS traces where more recently uploaded GPS traces are placed on top of the list. Transportation mode, date, total traveled distance, time duration, and average speed are shown for each GPS trace in the list. Two time series on the user page (Figure 4(a)) show the total traveled distance for the user and the average of all users in different years, so the user can compare his/her traveled distance with other users. Users can delete their GPS traces or upload new GPS traces in this page. For each new GPS trace, the user must determine the transportation mode through a drop-down menu.

The trace manager page (Figure 4(b)) shows a list of all GPS traces which have been assigned to this manager along with their summary statistics and the owner’s username. The time series on this page show the total traveled distance over all users for different transportation modes. The user manager page (Figure 4(c)) shows a list of all users who have been assigned to this manager. A bar chart on this page shows the total traveled distance for each user.

3. Data
OSM GPS traces cannot be used to populate the proposed database in this work due to their lack of transportation mode. GeoLife project, conducted by Microsoft Research Asia, collected GPS traces from 182 users between April 2007 and August 2012 (Zheng, Chen, et al., 2010; Zheng, Li, Chen, & Xie, 2008; Zheng, Liu, Wang, & Xie, 2008). The data is available for download on their website (Microsoft Research, 2012). The data includes 17,621 GPS traces with a total distance of 1.2 million km and a total duration of over 48,000 h. However, only a small portion of this data, shown in Figure 5, are associated with a transportation mode (Microsoft Research, 2010). Since transportation mode is an integral part of our database and applications, only this small portion of GPS traces qualifies for our work, which belongs to 32 users. All the 32 users and their associated GPS traces are used in this work. In the GeoLife data-set, transportation modes are stored in a separate plain text file for each user and include walk, run, bike, car, driving meet congestion, motorcycle, taxi, bus, subway, train, railway, plane, airplane, and boat. The transportation mode file includes start time and end time for each transportation mode. Therefore our first task was to associate the transportation modes to GPS traces which was challenged by multimodal GPS traces, occasional long gaps between sequential points, inconsistencies in timestamps (the time goes back), and implausible traveled distances in short times. We considered 20 min as a threshold for the gap between two sequential GPS points to decide when to split a GPS trace into two GPS traces. GPS traces are also split whenever the
transportation mode changes to create unimodal GPS traces. GPS traces with inconsistent times-
tamps or out of range longitudes and latitudes were excluded. Tables 1 and 2 show statistics of the
refined data.

The statistics listed in these tables, such as average speed, can be used as knowledge for other
applications. For example, average walking speed for a person can be used to personalize pedestrian
navigation systems or to compare people’s mobility behaviors in different regions. Some errors in
GPS traces are semantic and thus, cannot be easily detected. For example, the average speed for
one of the running GPS traces is 26.87 m/s, while the top running speed for people usually ranges
from 6.2 m/s to 11.1 m/s (Weyand, Sternlight, Bellizzi, & Wright, 2000). The knowledge gained from
these tables can help to detect such semantic inconsistencies.

| User name | Number of GPS traces | Number of Points | Total length (km) | Total duration (hour) | Average speed (m/s) |
|-----------|----------------------|------------------|-------------------|----------------------|---------------------|
| 1         | 446                  | 106,447          | 2,048             | 131.42               | 4.33                |
| 2         | 335                  | 193,595          | 5,809             | 206.19               | 7.83                |
| 3         | 23                   | 304              | 488               | 9.32                 | 14.54               |
| 4         | 34                   | 1,661            | 108               | 10.35                | 2.90                |
| 5         | 43                   | 9,709            | 166               | 13.23                | 3.49                |
| 6         | 424                  | 338,864          | 3,830             | 206.56               | 5.15                |
| 7         | 27                   | 4,573            | 195               | 17.70                | 3.06                |
| 8         | 1                    | 544              | 4                 | 0.30                 | 3.70                |
| 9         | 60                   | 35,840           | 826               | 24.01                | 9.36                |
| 10        | 126                  | 62,215           | 865               | 46.25                | 5.20                |
| 11        | 461                  | 322,395          | 3,731             | 276.45               | 3.75                |
| 12        | 1,221                | 524,080          | 6,627             | 407.39               | 4.52                |
| 13        | 178                  | 5,191            | 416               | 25.48                | 4.54                |
| 14        | 125                  | 56,889           | 302               | 31.67                | 2.65                |
| 15        | 18                   | 2,640            | 183               | 10.11                | 5.03                |
| 16        | 232                  | 74,849           | 1,026             | 93.21                | 3.06                |
| 17        | 154                  | 115,094          | 4,557             | 132.78               | 9.53                |
| 18        | 38                   | 22,951           | 523               | 29.30                | 4.96                |
| 19        | 140                  | 67,853           | 1,134             | 71.43                | 4.41                |
| 20        | 74                   | 50,306           | 358               | 33.31                | 2.99                |
| 21        | 475                  | 344,187          | 4,233             | 242.44               | 4.85                |
| 22        | 967                  | 475,492          | 34,047            | 737.50               | 12.82               |
| 23        | 20                   | 4,384            | 167               | 11.05                | 4.20                |
| 24        | 19                   | 387              | 58                | 4.39                 | 3.67                |
| 25        | 27                   | 1,691            | 65                | 11.45                | 1.58                |
| 26        | 46                   | 1,323            | 237               | 18.33                | 3.59                |
| 27        | 1,119                | 348,466          | 5,315             | 290.62               | 5.08                |
| 28        | 48                   | 2,587            | 143               | 17.14                | 2.32                |
| 29        | 37                   | 802              | 403               | 6.28                 | 17.83               |
| 30        | 162                  | 95,034           | 4,373             | 111.23               | 10.92               |
| 31        | 904                  | 584,153          | 4,934             | 335.31               | 4.09                |
| 32        | 527                  | 238,172          | 1883              | 144.15               | 3.63                |
| Overall   | 8,511                | 4,092,678        | 89,053            | 3,706.33             | 6.67                |
4. Detecting POIs

One of the popular applications of GPS traces, when they are available in large amounts, is finding POIs of people (Lane, Lymberopoulos, Zhao, & Campbell, 2010; Lian & Xie, 2011; Shaw, Shea, Sinha, & Hogue, 2013). These POIs can be places where a person lives, works, shops, or spends time. Knowing users’ POIs helps to provide them with more intelligent location-based or location-aware services. Following are the steps used to detect a specific user’s POIs:

(a) Since we define POIs as points that a user usually visits for resting, relaxing, working, shopping, exercising, or similar activities, POIs must either be the origin or destination of a trip. Therefore, call the first and last GPS points in each GPS trace significant points and discard the rest.

(b) For each significant point, count the number of other significant points falling in a circle of 50 m radius around it. Call this value the significance rate. Therefore, each significant point is associated with a significance rate.

(c) If two significant points are closer than 100 m to each other, discard the one with the lower significance rate. This is because their 50 m radius circles overlap and significant points falling in the overlap area are being counted for both of them. On the other side, having two POIs closer than 100 m to each other is not realistic.

Since the user with the id 12 has the largest number of GPS traces, we use his/her GPS traces to detect his/her POIs. Figures 6-10 show his/her POIs and their significance rates for different transportation modes.

For driving, walking, running, and bus GPS traces, the POIs are mostly focused in two very small areas, one in the north (referred to as POI\textsubscript{N}) and one in the south (referred to as POI\textsubscript{S}) in Figures 6-9. These are most probably his/her working and living places. For biking GPS races in Figure 10, on the other hand, the POIs are more spread over the city. A closer look at the biking POIs shows that the most significant one (at 40°5′, 116°20′) falls in the POI\textsubscript{N} area, though none of them falls in the POI\textsubscript{S} area. This can be used as an argument that the POI\textsubscript{N} area is the living place rather than the working place. Additionally, the majority of walking and running POIs also fall in the POI\textsubscript{N} area. Overlapping these POIs with a land-use map can reveal the name of locations and buildings.
5. Associating patterns in walking GPS traces with people's identities

Assume we have a walking GPS trace whose owner is unknown. In this section, we explore the possibility of finding its owner assuming he/she is among the system's users. GPS trace recognition, like face recognition, has applications in information systems and services. For example, Facebook can suggest you to tag your friends using their GPS traces like it does using their pictures.

This is a classification problem where classes are users of our system. Each walking GPS trace is a sample or observation whose feature vector includes: average speed, sampling rate, average longitude, and average latitude. To justify the selection of these four features for our classification...
problem, we investigate their distribution across classes using box plots, their linear independence from each other using correlation coefficients, and eigenvalues of the features’ covariance matrix. Figure 11 shows the box plot of sampling rate of walking GPS traces in each of the 32 classes. Sampling rate is one of the four features. Each person is a class in this plot, represented by an individual box. The box plot is a standardized way of displaying the distribution of data based on the five number summary: minimum, first quartile, median, third quartile, and maximum. More overlap
among boxes means that feature is less diverse and less helpful in distinguishing among classes. The box plots in Figure 11 show that the values of this feature are well diversified across different classes (little overlap among boxes) and they can be effective in recognizing classes. Figures 12–14 represent the same type of plot for the other three features: speed, longitude, and latitude. Little overlap among boxes, observed in these plots, similarly indicates their effectiveness in distinguishing among classes. Table 3 reports the correlation coefficient for pairs of features over all classes. If two features are strongly correlated (indicated by correlation coefficients close to ±1), there is not much sense in considering both of them in the feature vector. However, all correlation coefficients in Table 3 are close to zero. The almost equal eigenvalues of the features’ covariance matrix (1.1, 1.0, 1.0, and 0.9) also indicate equal significance of different features in our classification problem.

Bayesian classifiers transcend other classifiers in terms of minimizing the error probability which comes with the cost of their need to large amounts of training data (~10^l samples for each class where l is the number of features which is 4 in our case) in order to detect the underlying probability

---

Figure 10. POIs over biking GPS traces.

Figure 11. Box plot for sampling rate of walking GPS traces for different people.
Figure 12. Box plot for speed of walking GPS traces for different people.

Figure 13. Box plot for longitude of walking GPS traces for different people.

Figure 14. Box plot for latitude of walking GPS traces for different people.
density function (PDF) of features in each class (Theodoridis & Koutroumbas, 2009). Among different versions of Bayesian classifier, the Naïve Bayesian (NB) classifier is the right choice in this application because of the validity of the assumption that features in each class are independent. Independence of features in our classification problem is both conceptually sensible and quantitatively shown in Table 3. With this assumption, the NB classifier reduces the size of the required training data-set from $10^l$ to approximately $10 \times l$ (Theodoridis & Koutroumbas, 2009). Most of our classes qualify this training sample size requirement. An unlabeled walking GPS trace ($x$) is assigned to the class ($\omega_i$) with the largest posterior $p(\omega_i|x)$. To avoid presumptions about the overall shape of the PDFs, the non-parametric Parzen Window approach with a Gaussian Kernel (Theodoridis & Koutroumbas, 2009) is applied to estimate likelihoods $p(x|\omega_i)$. Two scenarios are considered for priors $p(\omega_i)$: assumed equal for all classes and estimated as the relative frequency of classes ($p(\omega_i) = n_i/n$ where $n_i$ is the number of training samples in class $\omega_i$ and $n$ is the total number of training samples).

Least squares (LS) and linear support vector machine (LSVM) among linear classifiers and non-linear support vector machine (NLSVM) with a Gaussian kernel (with $\sigma = 1$) among non-linear classifiers are selected here for experimental purposes. The smoothing parameter ($C$) is considered 10 and 325 for LSVM and NLSVM, respectively. These are the experimentally optimized values for $C$ as represented in Figure 15. Because LS, LSVM, and NLSVM classifiers can distinguish between only two classes, a separate classifier is trained for each pair of classes (one-to-one approach). Since there are 31 classes (the user with Id 8 in Table 1 is dismissed since he/she only has one GPS trace), there could be $C_2^{31} = 465$ pair of classes. Therefore, for each of the three aforementioned classifiers (LS, LSVM, and NLSVM), 465 pairwise classifiers need to be trained and the class with the largest number of wins in pairwise comparisons wins the unlabeled sample.

| Table 3. Correlation coefficient between pairs of features |
|----------------------------------------------------------|
| Sampling rate vs. longitude | Sampling rate vs. latitude | Sampling rate vs. speed | Longitude vs. latitude | Longitude vs. speed | Latitude vs. speed |
|----------------------------|-----------------------------|-------------------------|------------------------|-------------------|-------------------|
| 0.001                      | -0.088                      | 0.009                   | 0.039                  | 0.006             | -0.034            |

Figure 15. The overall accuracy of LSVM and NLSVM vs. the smoothing parameter ($C$).

Figure 16. The overall accuracy of different classifiers.
We also define the combined classifier by combining the results of all previous classifiers using the product rule. In the product rule, the combined posterior of a class is the multiplication of the posteriors obtained for that class from different classifiers (Theodoridis & Koutroumbas, 2009). A random classifier, which randomly assigns a GPS trace to one of the 31 classes, is also considered to depict the lowest expected accuracy of any non-random classifier.

The leave-one-out cross validation method (Theodoridis & Koutroumbas, 2009) is used to evaluate the generalization accuracy of different classifiers. Figure 16 shows the overall accuracy of different classifiers, where RC stands for random classifier, LSVM stands for linear SVM, LS stands for least squares, NLSVM stands for non-linear SVM, NB1 is the naïve Bayesian with equal priors, NB2 is the naïve Bayesian with relative frequencies as priors, and CC stands for the combined classifier. Figure 17 shows the sample size vs. recall and precision (obtained from NB2 classifier which revealed the best accuracy in Figure 16) for different classes.

As shown in Figure 17, having more or less training samples from a class does not necessarily reflect in higher or lower precision and recall for that class when it comes to Bayesian classifiers. Precision and recall for a class increases if that person’s walking characteristics follow a specific and discriminable pattern. More specific and different from others that pattern is, higher the recall and precision would be for that class. Therefore, choosing which characteristics of walking GPS traces need to be used as predictors for classification is quintessential because it is the combination of these characteristics which is supposed to reveal the specific and distinguishable walking pattern of each person.

6. Conclusions and future directions
Gathering GeoLife GPS traces in a relational database was the most cumbersome part of this work. However, the created database facilitated the access, analysis, and management of GPS traces. The metadata provided for the GPS traces in our database were used to produce time series and bar charts on the front-end, semantically cross check the accuracy of GPS traces, and detect useful human mobility patterns. We showed how our database can be used to detect users’ POIs and walking patterns, although there are many other such applications for this database, e.g. transportation network construction and traffic management. The results endorsed the database’s effectiveness in both applications.

Users need to determine the transportation mode while uploading their GPS traces in our system, which causes difficulties when a GPS trace is multimodal. More sophisticated approaches are required to collect GPS traces from users’ mobile devices directly and detect the transportation mode automatically or semi-automatically with the user’s help. The Naïve Bayesian classifier with longitude,
latitude, speed, and sampling rate as predictors achieved 47% overall accuracy in finding the owner of a walking GPS trace among 31 people (the accuracy of a random classifier is 3%). Investigating how much each predictor contributes in the overall accuracy and finding other predictors (e.g., time) which can boost this accuracy are also among our future research directions. Another important issue is protecting the users’ privacy which can be accomplished by anonymization (as applied in this study) or mixing GPS traces from different users. Privacy can be protected at different levels with stricter rules for people who are more concerned about their privacy and looser rules for those who do not mind revelation of specific aspects of their mobility patterns (Hashemi & Malek, 2012).

### Funding
The author received no direct funding for this research.

### Author details
Mahdi Hashemi1
E-mail: m.hashemi1987@gmail.com
1 School of Computing and Information, University of Pittsburgh, 135 North Bellefield Avenue, Pittsburgh, PA 15260, USA.

### Citation information
Cite this article as: Intelligent GPS trace management for human mobility pattern detection, Mahdi Hashemi, Cogent Engineering (2017), 4: 1390813.

### Cover image
Source: Author.

### References
Azevedo, T. S., Bezerra, R. L., Campos, C. A., & Moraes, L. F. (2009). An analysis of human mobility using real traces. In Wireless Communications and Networking Conference (pp. 2390–2395). Piscataway, NJ: IEEE.

Becker, R., Cáceres, R., Hanson, K., Isaacman, S., Loh, J. M., Martonosi, M., ... Volinsky, C. (2013). Human mobility characterization from cellular network data. Communications of the ACM, 56(1), 74–82. https://doi.org/10.1145/2398356

Bhattacharya, T., Kulk, L., & Bailey, J. (2015). Automatically recognizing places of interest from unreliable GPS data using spatio-temporal density estimation and line intersections. Pervasive and Mobile Computing, 19(1), 86–111. https://doi.org/10.1016/j.pmcj.2014.08.003

Chang, Y.-J., Liu, H.-H., Chou, L.-D., Chen, Y.-W., & Shin, H.-Y. (2007). A general architecture of mobile social network services. In Proceedings of the Convergence Information Technology Conference (pp. 151–156). IEEE.

Ferrari, L., Rosi, A., Mamei, M., & Zambonelli, F. (2011). Extracting urban patterns from location-based social networks. In Proceedings of the 3rd SIGSPATIAL International Workshop on Location-Based Social Networks (pp. 9–16). ACM.

Gonzalez, M. C., Hidalgo, C. A., & Barabási, A.-L. (2008). Understanding individual human mobility patterns. Nature, 453(7196), 779–782. https://doi.org/10.1038/nature06958

Harder, H., Nes, A. V., Jensen, A. S., Reinau, K. H., & Weber, M. (2012). Time use and movement behaviour of young people in cities: The application of GPS tracking in tracing movement pattern of young people for a week in Aalborg. In M. Greene, J. Reyes, & A. Castro (Eds.), Proceedings of the 8th International Space Syntax Symposium (p. 5538). Santiago de Chile: PUC.

Hashemi, M. (2017a). Reusability of the output of map-matching algorithms across space and time through machine learning. IEEE Transactions on Intelligent Transportation Systems. doi:10.1109/TITS.2017.2669085

Hashemi, M. (2017b). A testbed for evaluating network construction algorithms from GPS traces. Computers, Environment and Urban Systems, 66, 96–109. https://doi.org/10.1016/j.compenvurbsys.2017.08.003

Hashemi, M., & Korimi, H. A. (2014). A critical review of real-time map-matching algorithms: Current issues and future directions. Computers, Environment and Urban Systems, 48, 153–165. https://doi.org/10.1016/j.compenvurbsys.2014.07.009

Hashemi, M., & Karimi, H. A. (2016a). A machine learning approach to improve the accuracy of GPS-based map-matching algorithms. In Proceedings of the IEEE 17th International Conference on Information Reuse and Integration (pp. 77–86). Pittsburgh, PA: IEEE.

Hashemi, M., & Karimi, H. A. (2016b). A weight-based map-matching algorithm for vehicle navigation in complex urban networks. Journal of Intelligent Transportation Systems, 20(6), 573–590. https://doi.org/10.1080/15472765.2016.1166058

Hashemi, M., & Karimi, H. A. (2017). Collaborative personalized multi-criteria wayfinding for wheelchair users in outdoors. Transactions in GIS, 21(4), 782–795. https://doi.org/10.1111/tgis.2017.21.issue-4

Hashemi, M., & Malek, M. R. (2012). Protecting location privacy in mobile geoservices using fuzzy inference systems. Computers, Environment and Urban Systems, 36(4), 311–320. https://doi.org/10.1016/j.compenvurbsys.2011.12.002

Lone, N. D., Lymberopoulos, D., Zhao, F., & Campbell, A. T. (2010). Hapori: Context-based local search for mobile phones using community behavioral modeling and similarity. In Proceedings of the 12th ACM International Conference on Ubiquitous Computing (pp. 109–118). ACM. https://doi.org/10.1145/1864349

Lee, K., Hong, S., Kim, S. J., Rhee, I., & Chong, S. (2009). Slaw: A new mobility model for human walks. In 28th Annual Joint Conference of the IEEE Computer and Communications Societies (pp. 855–863). Rio de Janeiro: IEEE.

Li, X., Pan, G., Wu, Z., Qi, G., Li, S., Zhang, D., ... Wang, Z. (2012). Prediction of urban human mobility using large-scale taxi traces and its applications. Frontiers of Computer Science, 6(1), 111–121.

Lian, D., & Xie, X. (2011). Learning location naming from user check-in histories. In Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (pp. 112–121). ACM.

Liao, L., Patterson, D. J., Fox, D., & Kautz, H. (2007). Learning and inferring transportation routines. Artificial Intelligence, 171(5–6), 311–331. https://doi.org/10.1016/j.artint.2007.01.006

Liu, L., Andris, C., & Ratti, C. (2010). Uncovering cabdrivers’ behavior patterns from their digital traces. Computers, Environment and Urban Systems, 34(6), 541–548. https://doi.org/10.1016/j.compenvurbsys.2010.07.004

Liu, X., & Karimi, H. A. (2006). Location awareness through trajectory prediction. Computers, Environment and Urban Systems, 30(6), 741–756. https://doi.org/10.1016/j.compenvurbsys.2006.02.007

Microsoft Research. (2010, November 24). GPS trajectories with transportation mode labels. Retrieved November 18, 2015, from GeoLife: Building social networks using human location history http://research.microsoft.com/apps/pubs/?id=141896
Microsoft Research. (2012, August 9). Geolife GPS trajectories. Retrieved November 18, 2015, from GeoLife: Building social networks using human location history http://research.microsoft.com/en-us/downloads/b16d3559d-a164-4e9e-9fd4-d2a38f2b2e13/

OpenStreetMap. (n.d.). OpenStreetMap. Retrieved November 18, 2015, from OpenStreetMap traces https://www.openstreetmap.org/traces

Pan, G., Qi, G., Zhang, W., Li, S., Wu, Z., & Yang, L. T. (2013). Trace collection and mining for smart cities: Issues, methods, and applications. IEEE Communications Magazine, 51, 120–126. https://doi.org/10.1109/MCOM.2013.6525694

Patel, K., Chen, M. Y., Smith, I., & Landay, J. A. (2006). Personalizing routes. In Proceedings of the 19th annual ACM symposium on User interface software and technology (pp. 187–190). ACM.

Patterson, D. J., Liao, L., Fox, D., & Kautz, H. (2003). Inferring high-level behavior from low-level sensors. In Proceedings of the Ubiquitous Computing Conference (pp. 73–89). Springer.

Peng, C., Jin, X., Wong, K.-C., Shi, M., & Li, P. (2012). Collective human mobility pattern from taxi trips in urban area. PLoS ONE, 7(4), e34487. https://doi.org/10.1371/journal.pone.0034487

Phithakkitnukoon, S. T., Lorenzo, G. D., Shibasaki, R., & Ratti, C. (2010). Activity-aware map: Identifying human daily activity pattern using mobile phone data. In A. A. Salah, T. Gevers, N. Sebe, & A. Vinciarelli (Eds.), Human behavior understanding (pp. 14–25). Berlin Heidelberg: Springer.

Shaw, B., Shea, J., Sinha, S., & Hogue, A. (2013). Learning to rank for spatiotemporal search. In Proceedings of the 6th ACM international conference on web search and data mining (pp. 717–726). ACM.

Song, C., Qu, Z., Blumm, N., & Barabasi, A.-L. (2010). Limits of predictability in human mobility. Science, 327(5968), 1018–1021. https://doi.org/10.1126/science.1177170

Theodoridis, S., & Koutroumbas, K. (2009). Pattern recognition (4th ed.). Elsevier.

Weyand, P. G., Sterlign, D. B., Bellizzzi, M. J., & Wright, S. (2000). Faster top running speeds are achieved with greater ground forces not more rapid leg movements. Journal of Applied Physiology, 89(5), 1991–1999.

Wood, H. (2013, January 6). 1 Million OpenStreetMappers. Retrieved November 18, 2015, from OpenStreetMap Blog https://blog.openstreetmap.org/2013/01/06/1-million-openstreetmappers/

Yu, X., Pan, A., Tang, L.-A., Li, Z., & Han, J. (2011). Geo-friends recommendation in GPS-based cyber-physical social network. In Proceedings of the Advances in Social Networks Analysis and Mining Conference (pp. 361–368). IEEE.

Zheng, V. W., Cao, B., Zheng, Y., Xie, X., & Yang, Q. (2010). Collaborative filtering meets mobile recommendation: A user-centered approach. Proceedings of the 24th AAAI Conference on Artificial Intelligence, 10 236–241.

Zheng, Y., Chen, Y., Li, Q., Xie, X., & Mao, W.-Y. (2010). Understanding transportation modes based on GPS data for web applications. ACM Transaction on the Web, 4(1), 1–36.

Zheng, Y., Li, Q., Chen, Y., & Xie, X. (2008). Understanding mobility based on GPS data. In Proceedings of ACM conference on Ubiquitous Computing (pp. 312–321). Seoul: ACM.

Zheng, X., Liu, L., Wang, L., & Xie, X. (2008). Learning transportation modes from raw GPS data for geographic application on the web. In Proceedings of International Conference on World Wide Web (pp. 247–256). Beijing: ACM.