Using Spatio-temporal Trajectories to Monitor Construction Sites for Safety Management

ABSTRACT

Construction sites are one of the most hazardous working environments due to their dynamic and decentralized nature. Despite numerous efforts to reduce fatalities and risks associated with working on construction sites, such accidents continue to occur. The advancements in location acquisition and mobile communication techniques offer various means to monitor construction sites with the help of mobility data. Information derived from such data offers tremendous opportunities to track and analyze the mobility of moving objects, that can contributes to improve safety management processes. However, the working of such systems and processes requires preprocessed positioning data of moving objects as raw data captured from location acquisition devices such as GPS device is generally available as discrete points and do not hold enough information to understand the mobility. In this article, we have conducted a review of applications that use mobility data also known as trajectories. After a detailed literature review, an application of trajectories carrying spatio-temporal information for worker safety on construction sites is discussed. The proposed application collects raw GPS data and reconstructs workers` trajectories using different data processing algorithms. Once the trajectory data is cleaned and processed, it can be further enriched with semantic and contextual information to enable the desired interpretation of workers` movements on construction sites and ultimately attempts to improve work zone safety by reducing fatalities. However, the scope of this paper is kept limited to data processing of GPS trajectories.

Keywords

Health and Safety (H&S); construction sites; fatal accidents; mobility

1. INTRODUCTION

International occupational health and safety statistics shows that the construction industry experiences one of the highest accident rates of all industries. According to the National Census of Fatal Occupational Injuries conducted by U.S. Bureau of Labor Statistics in 2015, out of 4,836 fatal work injuries 19 % of fatalities were recorded from the construction industry. The major reasons of fatalities were related to the unsafe human behaviors, difficult site conditions, workers falling from heights and striking against or being struck by moving objects. Despite numerous efforts to reduce fatalities on construction sites, such accidents continue to occur as shown is Figure 1.

The advancements in location acquisition and mobile communication techniques based on GSM (Global System for Mobile Communications), Bluetooth, Wi-Fi and other wireless sensing technologies offers various opportunities to collect mobility data from construction sites for safety management. With the help of GPS (Global Positioning System) equipped devices and vehicles, Radio Frequency Identification (RFID) tag tracking and location-aware wireless sensors, the possibilities to collect larger mobility data is increased significantly [4]. These technologies can generate huge amount of data that contains time varying geographic positions of moving objects called trajectories. Such trajectories captured from moving objects have unique

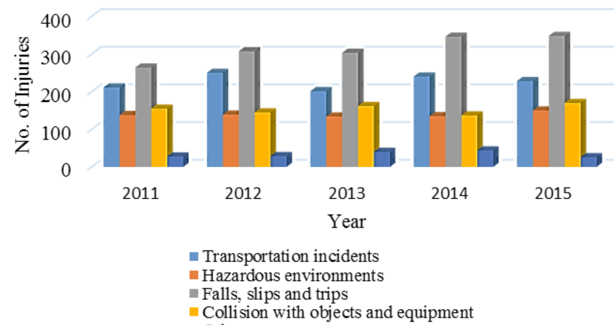

Figure 1. Fatal injuries in U.S. construction industry [OSHA]

features; they are in the form of streaming data which means that data will keep on updating over time and includes spatial information [5]. Trajectories with such information are known as spatio-temporal trajectories. In recent years, it is technically more convenient and economically cheap to acquire the spatio-temporal trajectories of moving objects from a wider region in less time [6]. The analysis of such trajectories can lead to establish better understanding and knowledge discovery of moving objects [7].

As the data volume is growing, it demands the need to improve the existing capabilities of managing and analyzing trajectory data [8, 9]. A lot of research has been established ranging from trajectory data models and query languages to deployment aspects, such as indexing and query optimization methods. Moving Object Database (MOD) and Spatio-Temporal database are the some of the examples of trajectory databases [10]. The objective of such databases is to extend database technology in order to support the representation and querying of trajectory data. Their primary research objectives are similar to traditional databases such as creating ad-hoc data representations, indexing, storage and querying. However, raw trajectory data should be first cleaned and preprocessed before moving it towards data management [11]. Datasets acquired by GPS devices are often suffer from noise [12]. Therefore, cleaning raw trajectories cannot be ignored when recreating meaningful trajectory episodes from it. The trajectories of moving objects are continuous in nature in real world scenarios [13]. Because of data collection and storage devices limitations, continuous movements of moving objects are collected and stored in a discrete format such as a sequence of GPS points(x, y, t). As tracking time of GPS device increases, the trajectory data in applications will gradually grows. Such huge amount of data can sooner or later causes storage challenges. Therefore, trajectory data reduction should be done as a pre-processing step before trajectory data management [14]. The primary aim of trajectory reduction is to decrease the computational complexity and reducing the time to query the trajectory data [14]. Once trajectory data is cleaned, processed and reduced, it is now ready to be saved in a trajectory database. Many applications in the mobility area require semantic interpretation of mobility information as well, that is not possible to extract from raw GPS trajectories. As the physical trajectories can easily be recorded using GPS devices but the semantic interpretation of the mobility data is still a big research challenge [15].

The paper is organized as follows: in section 2, trajectories and their movement related characteristics are described. In section 3, applications of spatio-temporal trajectories are presented. Section 4, is based on the preprocessing of raw trajectories using different algorithms for worker safety on a construction site application. Section 5 presents the discussion of the presented work and conclusion is discussed in section 6.

2. Trajectories and their Movement-Related Characteristics

A trajectory is the path that is followed by a moving object in space as a function of time [16]. It is acquired as a series of location points having time stamps, denoted as follow;

$$\{\langle x_1, y_1, t_1\rangle, \langle x_2, y_2, t_2\rangle, \dots, \langle x_N, y_N, t_N\rangle\}$$

Here, x_i and y_i represent x and y geographical coordinates at time t_i, N is the total number of points in the trajectory data and it is usually followed by a sampling interval. There are many approaches to collect trajectory data. It can be based on the change in time, location, event or various combination of these approaches. As trajectory data holds multi-faceted characteristics, these are used to analyze and understand mobility [17]. Mobility related characteristics include:

1) Time (i.e., position of object mobility on the timescale)
2) Position of the object in geographical coordinate system
3) Direction of the object
4) Speed of the object
5) Change in direction
6) Acceleration (i.e., change in speed)
7) Distance travelled

These characteristics can directly be computed from the raw trajectories and are well-suited for the applications that are performing localization of objects in motion. For example, where was the position of John at 10am on June 20, 2017? However, most decision making applications requires additional information with the trajectory data from the application context. Other than location data points of trajectories, there exists an important characteristic of mobility data that is the variety of various travel means by which a mobility has taken place [18]. This is of special interest in domains such as urban planning as selection of location and travel mean may be interconnected. For example, to analyze trajectories of a person going from home to office in a city, it is necessary to have some information about the city and its infrastructure such as information about shops and restaurants etc. Such data would help to visualize person going to office trajectories in more detail in terms of the point of interest locations rather than just by the geographical coordinates of the location. The process of supplementing the GPS trajectories with additional data is known as semantic enrichment process [19]. This additional data categorizing stops and moves and distinguishing different types of moves is known as annotation that is attached to a trajectory either to some of its parts or as a whole [20]. An annotation value is simply an attribute value that can be an "on-tram" or an "on-bus", a possible value for TransportationMeans annotation in case of a person going to an office scenario. An example of semantically enriched trajectory could be the following [20].

(Begin, home, 8am, -) → (move, road, 8am-8:30am, walk) → (move, road, 8:30am-9am, on-tram) → (stop, office, 9am-5pm, work) → (move, road, 5pm-5:30pm, on-tram) → (move, road, 5:30pm-6:00pm, walk) → (End, home, 6:00pm, -)

An above example includes generic movement characteristics (e.g.,stops and moves), application specific geographical objects (e.g., office and work) and also additional behavioral context (e.g., work). Annotations can be complex values as well, that may be collected from raw GPS data, contextual data or inferred by reasoning (giving a TransportationMeans by calculating a speed of a person) [20 - 22]. Once the trajectories are annotated and stored in a trajectory database, it can be used to detect and analyze trends and behaviors of moving objects.

3. Literature Review

Before using trajectory data in applications, it should be processed according to the applications` requirements. Trajectory data processing deals with the algorithms for trajectory reconstruction. For an initial preprocessing, research has been done to clean the data by removing the noise and outliers and then compressing it. Noise in trajectories is small distortions that have the potential to affect speed, acceleration and the travel distance estimates. However, their impact can be reduced by smoothing techniques [23]. Marketos et al. proposed an online approach to filter outliers in a trajectory data by taking the maximum allowed speed of the moving object as a reference point [24]. In addition, Jun et al. proposed a modified version of Kalman filter to control the outliers in GPS data with more effective technique [25]. Their algorithm demonstrates superior results as compared to other data smoothing techniques and it can be used for the real time data smoothing. There are also many works discuss the data compression methods of a trajectory data [26, 27]. For instance, Meratnia and de By proposed an algorithm that considers both temporal and spatial parameters in order to compress trajectories [28]. In their proposed technique, distance and speed thresholds have been used for data compression. Although, choosing proper thresholds are totally dependent on the application and having a clear understanding of an object in motion will help in the selection of appropriate thresholds. Their proposed algorithms can be used for offline as well as for online data reduction.

After data compression, segmentation is another important step to understand mobility data. Buchin et al. proposed a framework to segment a trajectory based on spatio-temporal criteria [29]. Segmentation is achieved in such a way that each generated segment is made homogenous and it should fulfill the spatio-temporal criteria. These criteria include location, speed, heading, velocity, curvature and shape. Under any on these criteria or any combination of these criteria, trajectory segmentation can be achieved in an efficient and optimal way. As an extension of Buchin et al. study [29], Maike et al. proposed a framework to segment a trajectory based on the individual movement states of a moving object [30]. Their approach requires expert knowledge and manual input from the user to describe movement states with spatiotemporal criteria. After processing of GPS trajectories data, it can now be used in variety of applications. As data acquired from GPS lacks in semantic information for the behavior analysis of a moving object. These GPS trajectories can further annotate using tools and techniques offered by data mining and machine learning concepts. However, it will raise the requirement for a training data. Trajectories resulting from annotation processes have been used in various applications such as urban indoor activity detection and wild life monitoring [31 - 33]. Such applications will help to understand the factors behind the movement of moving objects and these factors can be dependent on each other [34]. In [35], a prototype is developed for maritime surveillance system for better understanding the vessel trajectories. A major contribution of proposed system is to offer a reliable way to model semantic trajectories as well as semantic events. Geospatial trajectories have also been enriched with domain knowledge and contextual information to construct trajectory data-ware houses in order to support business intelligence processes in any organization [36, 37]. These data-ware houses are used to offer context-aware services such as recommender systems or digital assistance providing higher level of personalization to people for travelling [38, 39]. Once trajectories dataset is processed and maintained, hidden patterns can also be extracted [40]. Moreover, this will not only benefit in the exploration of important aspects of semantic trajectories but can also help to automate the processes of clustering the trajectories based on the stored information using machine learning techniques [41, 42]. Summary of these applications has been presented in a form of a Table 1.

4. Construction Workers Safety Monitoring Application

After an extensive review of applications, it is concluded that, spatio-temporal trajectories can also be used to track workers on construction sites. For this, a prototype system application is proposed to achieve the goal of visualizing trajectory data for safety management of workers. Building supervisor and H&S manager are the two roles have been identified from the literature for managing construction sites. These roles have been taken into account for the development of a prototype system as mentioned in Figure 2. The prototype system application focuses on the following:

- Capturing raw GPS data from construction workers` handheld devices on a constuction site.
- Aggregating and saving the collected GPS data to a cloud server having a centralized database configured.
- Transforming GPS data into raw trajectories.
- Executing preprocessing algorithms on a trajectory data to extract the movements of workers.
- Reducing the trajectory data before storing for future analysis, once preprocessing tasks are completed.
- Presenting trajectories visualizations to building supervisors and H&S manager for an effective construction site monitoring and ensuring safety at work.

4.1 Choice of Development Environment

The prototype system uses Mongodb for data storage and R platform for processing the GPS data. Mongodb server is used because of:

- Open source solution
- Cloud based technology
- Store massive datasets efficiently
- Easy to configure with a programming platform

R platform is used because of:

- Open source solution
- Offers integrations with JSON based databases such as Mongodb
- Offers packages to support geospatial data
- Provides libraries for visualizing massive datasets in high quality graphs

4.2 Prototype System Functionality

This section presents the functionality of a prototype system which is designed to monitor the geospatial trajectories of construction workers to minimize the safety hazards associated with working on a construction site. The proposed system can be divided into three main layers which are: GPS data acquisition, preprocessing the GPS data and generation of its visualizations. The data acquisition layer consists of smartphones that sends GPS data of workers at a defined interval. A single GPS data record consists of a user name, time stamp, latitude, longitude and level of floor. A screenshot of a GPS Common Seperated File (CSV) used for this application is shown below.

An application programming interface is designed to acquire GPS data from the workers` smart phone through wireless access points, aggregating and then storing it in a Mongdb server that is configured on a cloud. A data connection is established between a Mongodb and R studio to process the trajectories data. Mentioned below are the tasks that R studio will execute after retrieving data from a Mongodb server.

4.2.1 Data Cleaning

Data cleaning is the first step that R will perfom on a data retireved from a Mongodb. It is defined as a process of detecting and removing errors and irregularities from the data to improve the data quality. Real-life trajectories captured from GPS devices usually suffer from noise and to make data relaible for the application level of noise should be reduced. There can be various reasons for having noisy GPS trajectories, however sampling and measurement misadjustments, sensor battery outages and signal losses are some of them [43]. Due of these reasons, two types of errors occur in a GPS data: these are systematic and random. Systematic errors are occured due to low number of satellites available that's generates invalid GPS positions whereas, random errors are minor distortions from the true values occur due to atmospheric effects and receiver issues [43]. These both types of errors refer to the spatial positions. Whereas, temporal positioning is mostly accurate because of highly calibrated clocks operating in the satellites. For noise reduction, there exists mean and median filters which are the most simplied forms of filters for smoothing the GPS data. These filters are the types of causal filters as their dependency are only on the past values to compute the estimates. For this research, median filter is used because of its robustness characteristic whereas, mean filter is not recommended because it is highly sensitive to outliers [43]. In a median filter, for a measured point z_i, the estimate of the unknown value is the median of z_i and its $n-1$ predecessors in time. The median filter is based on sliding window mechanism covers n temporally adjacent values of z_i as shown in below mentioned equation.

$$\hat{x}_i = median\{z_{i-n+1}, z_{i-n+2}, z_{i-n+3}, \dots, z_{i-1}, z_i\}$$

Choice of a median filter for our data cleaning is made also because of high samplying rate of our GPS data; that makes median filter a good option. Though, if a sampling rate of GPS data is too low than a median filter is not recommended and advanced filters such as Kalman filter can be considered for noise reduction. Figure 3 shows two trajectories of two workers working on a construction site, whereas figure 4 shows actual and filtered trajectory of a worker A.

Table 1. Applications of Spatio-temporal trajectories

Use cases	**Detecting semantic outliers from moving objects [23].**	**Semantic trajectory compression for representing urban movements [26].**	**Compressing trajectory data by exploiting the semantic embedding of movement in a geographical context [27].**	**A geo-ontology design pattern for semantic trajectories [31].**	**Semantic middleware for trajectories to enable annotating trajectories [32].**	**Automated semantic trajectory annotation with indoor point-of-interest visits in urban areas [33].**	**Application of life trajectories for the modeling and analytics of human mobility in the urban context [34].**
Building indoor/outdoor	Outdoor	Outdoor	Outdoor	Outdoor	Outdoor	Indoor	Outdoor
Dataset	Dataset of school buses dataset consists of 145 trajectories of two school buses. The second dataset is of are trucks which consists of 276 trajectories of 50 trucks.	Everyday trajectories reflecting movement between regularly visited places of four persons.	Data contains geometric representation of the path contains 115 points in space-time (115 tuples of (x,y,t)). It further comprises 52 events, i.e., 52 intersections and stops along the way.	An individual's trajectory data recorded by a handheld GPS receiver and an animal tracking data retrieved from the MoveBank, an online database.	(1) 3 million GPS records of two Lausanne taxis, collected over 5 months (2) 2 million GPS records of 17,241 private cars tracked in Milan during one week; (3) a GPS trace of 2-hour drive of a private car in Seattle.	Data from 132 mobile devices containing the trajectory id and fields of latitude, longitude, altitude, accuracy, speed, and timestamp.	Data of fifty people to assess the model.
Findings	Extracted outliers from semantic trajectories to understand the unusual behaviors.	Proposed algorithm compresses trajectory data with a higher compression rate (89.53%) and drastically reduces the amount of data to be stored.	STC algorithm achieves a high compression rate. Instead of the 115 original points, it ends up with only 6 items, which corresponds to a compression rate of 94.78%.	Proposed design pattern is used to semantically annotate trajectory data of navigation and wildlife monitoring.	A framework is proposed to support semantic enrichment of trajectories exploiting both the geometric properties of the stream and the background geographic and application data.	Proposed an algorithm for the automated detection of visited points-of-interest. It extracts the actual visited points-of-interest well in terms of precision for the challenging urban indoor activity detection.	Proposed model allows a better understanding the reasons why and circumstances in which people are moving, whether they depend on external or on internal factors.
Key components/ technologies	Weka-STPM toolkit	-	GPS	Web Ontology Language (OWL)	Java 6 platform, PostgreSQL 8.4 with spatial extension PostGIS 1.5.1	PhoneGap platform	-

Use cases	Maritime surveillance system based on semantic events [35].	Semantic trajectories in Mobile Workforce Management (MWM) applications [36].	Mob-warehouse, a trajectory data warehouse to enrich trajectory data with domain knowledge [37].	Personalized RoutE Guidance System (PaRE) to provide higher-quality navigation directions [38].	Efficient frequent sequence mining on taxi trip records using road network shortcuts [39].	Extraction of hidden shared structure among human trajectories [40].	A semantic trajectory data model to define important aspects of semantic trajectories [41].	Discovering semantic places from GPS trajectories [42].
Building indoor/outdoor	Outdoor	Indoor and outdoor	Outdoor	Outdoor	Outdoor	Outdoor	Outdoor	Outdoor
Dataset	More than one million vessels positions with attributes such as the unique identification number of the vessel, the position, speed and rate of turn.	Data retrieved from the MWM system, the sensors of the mobile device and the GIS.	A trajectory dataset of people traveling by car in Milan (Italy), during one week. The dataset contains track of 16,946 cars and 48,906 trajectories for a total of 1,806,293 points.	Real trajectory dataset from Planet.gpx, which contains the GPS traces uploaded by OpenStreetMap users within 7.5 years.	17,558 taxi trip records in New York City over a month's period.	A collection of 230,000 GPS traces of taxi cabs in Beijing, China over a month's period.	A dataset of trajectories of tourists moving in Rome to visit the main city attractions.	Real GPS trajectories collected from 10 participants for nearly two months.
Findings	A prototype is developed to allow the maritime operator to better understand changes in the velocity of a vessel by analyzing semantic trajectories and semantic events.	Architecture is proposed to support business intelligence processes by collecting the information captured by the sensors of the mobile devices and analyzing and annotating the trajectory high-level activities using context information from the MWM system and geographic information providers.	A model is proposed, where the spatio-temporal component of trajectory data is properly integrated with context related information like transportation means, performed activities and mobility patterns.	Proposed system can reduce the number of navigation directions by more than 60% while still providing enough information for user to follow the route.	Results shows that runtimes of frequent sequence mining on shortcut sequences are orders of magnitude faster than on original road segment sequences.	Introduced the idea of "Pathlet Dictionary" to represent spatial regularities in a trajectory dataset and presented an eff ective algorithm to learn pathlet dictionaries from large collections of trajectories.	More semantics are added to raw trajectory data for real applications such as; tourism and animal behavior.	Hierarchical clustering algorithm is proposed to extract visit points from the GPS trajectories, and then these visit points can be clustered to form physical places.
Key components/ technologies	Triple store, Semantic Web Rules Language (SWRL), Spatio-temporal Inference engine	MWM system, OpenStreetMap and Android mobile devices	-	OpenStreetMap and Java platform	Spatial databases (PostgreSQL, PostGIS) and ArcGIS	-		

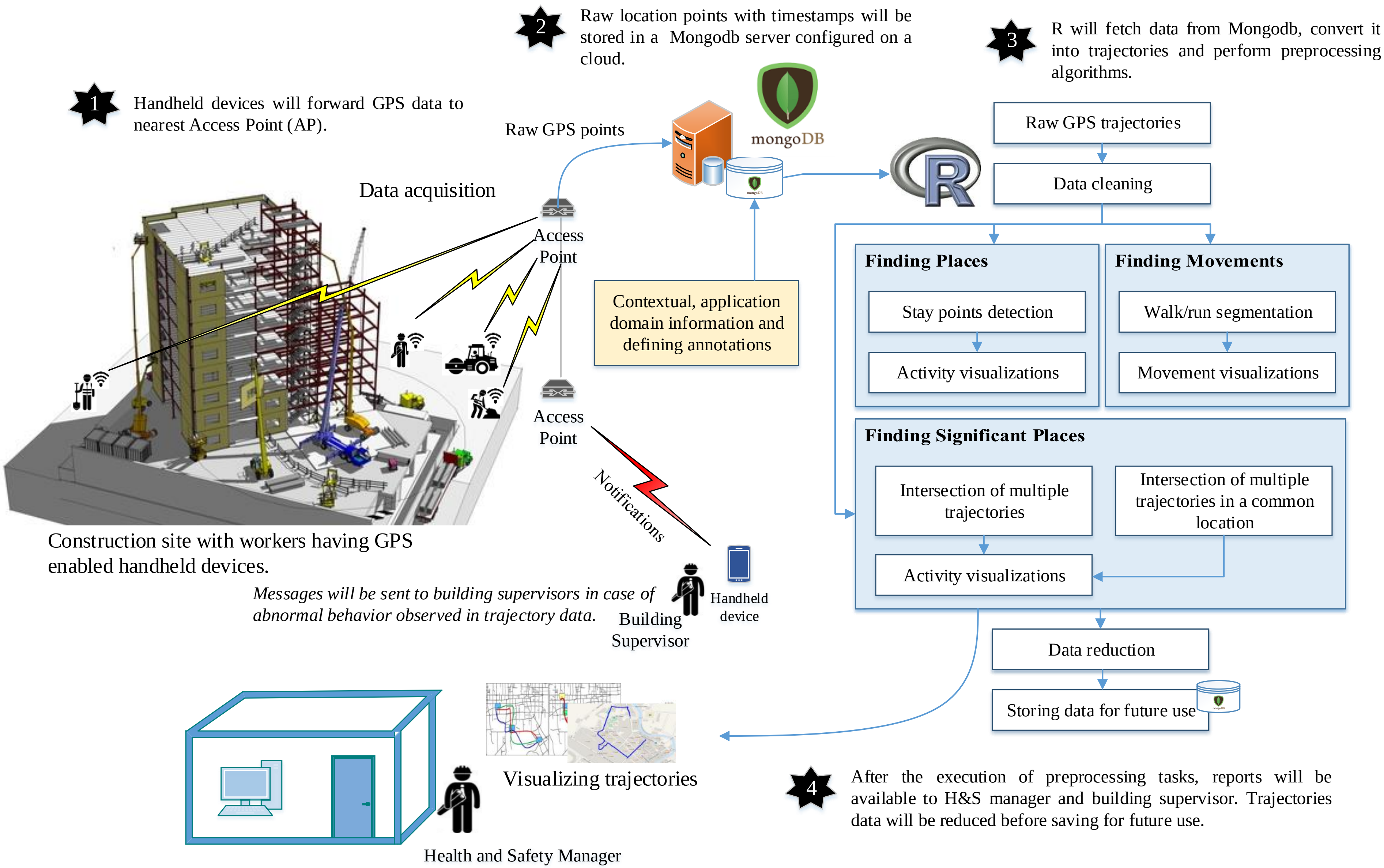

Figure 2. Preprocessing spatio-temporal trajectories for safety management on construction sites

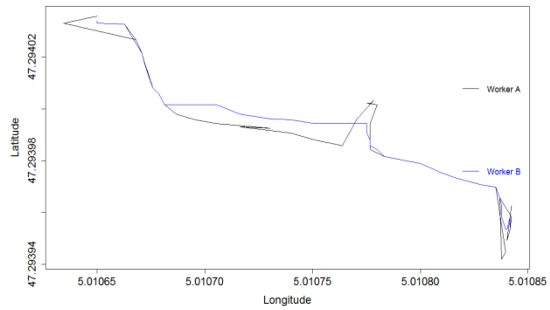

Figure 3. Trajectories of two different

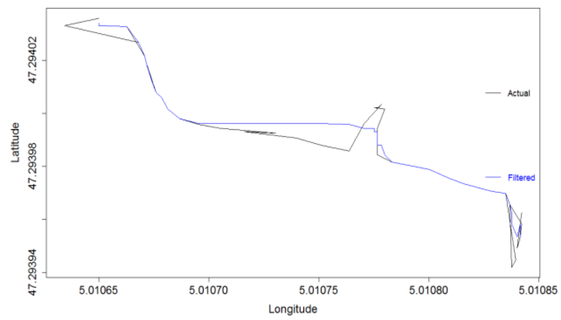

Figure 4. Actual and filtered trajectory of

4.2.2 *Stay Points Detection*

After cleaning the GPS trajectories, our preprocessing algorithm in the R environment will calculate stay points of workers as shown in Figure 5. Stay points are the geographic location points where an individual has spent a significant time within a certain distance. The calculation of stay points depends on two parameters, a distance threshold (D_{thresh}) and a time threshold (T_{thresh}). The values of these both parameters are application dependent and set manually in an algorithm. For our application, we need to find the locations on a contruction site where workers have spend more time than required. Calculating such stay points will help building supervisors and H&S manager to visualize the time spent on particular locations by the workers, monitoring their level of mobility and safety plans can be implemented accordingly. Li and Zheng et al. [43] proposed an algorithm for finding stay points in trajectories. A single stay point s can be treated as a virtual location point characterized by a set of successive GPS points $Z = \{z_m, z_{m+1}, z_{m+2}, \ldots, z_n\}$, $\forall\, m < i \leq n$, $Distance(z_m, z_i) \leq D_{thresh}$ and $|z_n.T - z_m.T| \geq T_{thresh}$. Formally, conditioned by Z, D_{thresh} and T_{thresh}, a stay point $s = (Latitude, Longitude, arrivaltime, leavingtime)$. Where,

$$s.latitude = \sum_{i=m}^{n} z_i.Latitude \Big/ |Z|$$

$$s.lontitude = \sum_{i=m}^{n} z_i.Longitude \Big/ |Z|$$

For an average latitude and longitude of the collection Z, $s.arrivaltime = z_m.T$ and $s.leavingtime = z_n.T$ represent a worker`s arrival and leaving times on a stay point (s). This algorithm first checks if the distance between a point under consideration and its successors in a trajectory larger than a specified threshold. Then is calculates the time interval between a point and the last successor that is within the distance threshold. If the time span is larger than a given threshold, a stay point is will be detected in a trajectory. For our application, there are four stay points have been detected in a worker A trajectory as shown in Figure below. However, number of stay points can be increased or decreased as these are totally dependent of the values of a distance threshold (D_{thresh}) and a time threshold (T_{thresh}).

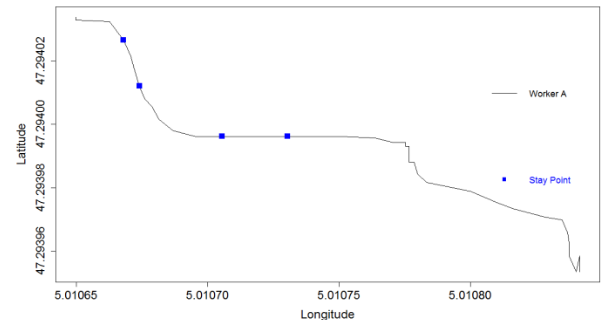

Figure 5. Stay points detection in a worker`s trajectory

4.2.3 *Segmentation of a trajectory*

For our application, we need to segment trajectories for finding types of movements which are being carried by the workers on a construction site. Segmentation is a process of dividing a trajectory into various segments to reduce the computational complexity but increasing the opportunities to mine richer knowledge. Segmentation can be done based on: time interval, shape of a trajectory and semantics meaning of location points in a trajectory. As constuction sites are very dynamic in nature and its very important to analysis the types of movements workers do on the site. Quick movements in workers trajectory data will be a red flag for building supervisors and H&S managers and highlights the abnornal situation occurred on a construction site.

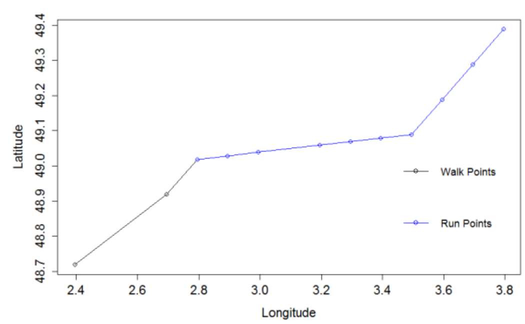

Figure 6. Segmentation of a worker`s trajectory based on speed

Based on our application requirement, we have used segmentation method that is based on semantics. Zheng et al. proposed a walk-based segmentation method [44, 45]. Their proposed algorithm calculates walk points and run points based on the point`s speed and acceleration. The trajectory can then be divided into alternate walk and run segments as shown in figure below.

4.2.4 *Intersection of two trajectories or Intersection between trajectories and a location*

It is important for our application to calculate the workers movements that have been carried collectively. Such understanding will depicits that how many workers are working on a same activity and how many will be affected on the occurance of any safety hazard on that particular location. As worker trajectory is a trace in the form of chronologically ordered points, distance from one worker trajectory to an other worker trajectory or a distance between workers trajectories and a common location can be found using a Haversine distance formula as mentioned below [46].

$$d = 2r\sin^{-1}\left(\sqrt{\sin^2\frac{\emptyset_i - \emptyset_j}{2} + \cos\emptyset_i \cos\emptyset_j \sin^2\frac{\varphi_i - \varphi_j}{2}}\right)$$

Where "r" is the Earth radius, $\emptyset$ and φ are the latitudes and longitudes of points "i" and "j". This formula is useful to find points of intersection in trajectories to visualize the esembly points within the area of interest.

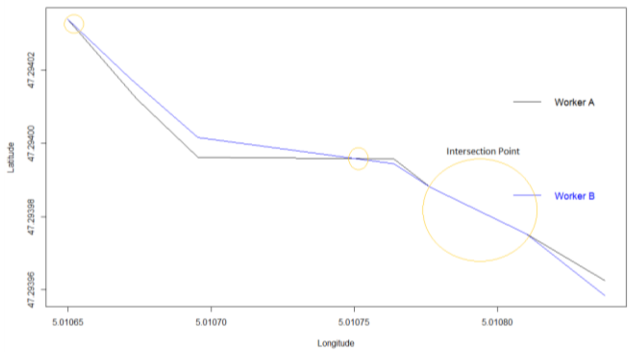

Figure 7. Points of intersection of two trajectories

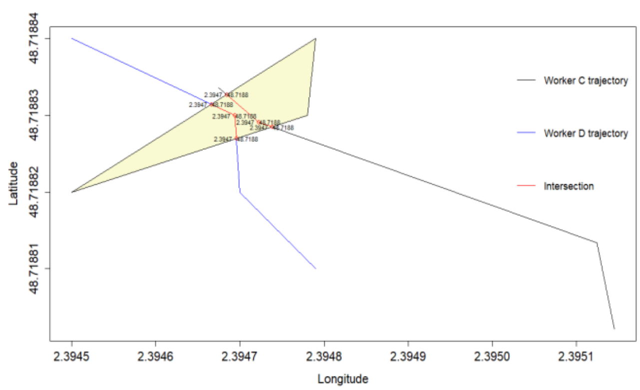

Figure 8. Points of intersection of two trajectories in a location

In addition, there can be some situations on construction sites where there is a need to find the closest point between two workers trajectories. For these scenarios, Haversine distance formula can also be used as shown in figure below.

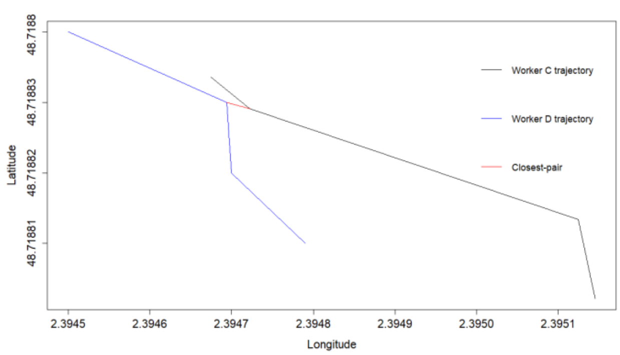

Figure 9. Calculating a closest point between two trajectories

4.2.5 *Trajectory data reduction*

Once cleaning and preprocessing of trajectory data is done, it is now required to store this data in a trajectory database for future use and analysis. Before storing, this data should be reduced in size. Potamias et. al discuss that a location point should be incorporated in a trajectory as long as it shows any change in the trajectory path [47]. As long as the location of an incoming GPS point can be predicted from previous points then this location point can be discarded safely as it will contribute very minute information. Batched compression and online data reduction are two methods to reduce a size of a trajectory [44]. Batch compression algorithms produce higher quality results when compared to online compression algorithms. For our application, we have used batched compression technique as our data is already been captured for a specific duration and its volume is kept limited for a deeper understanding. The most common form of batch compression algorithm for a data with higher sampling rate is the uniform sampling algorithm that is based on-line generalization mechanism. The main idea of line-generalization is to retain a fraction of the spatiotemporal data in the original trajectory without taking into account the redundancy of data points [43 - 45]. It keeps the every i-th data points (3rd, 6th, 9th, etc.) and discard the rest of points. As a results, in the future, reduced trajectories can be reconstructed that are an approximation of an original trajectories with a coarser granularity. Below figures shows the actual and reduced versions of a same trajectory using a uniform sampling algorithm.

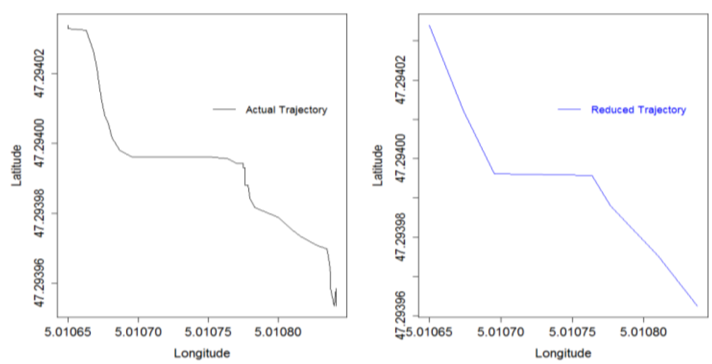

Figure 10. Actual and reduced worker trajectory

In addition to the uniform sampling algorithm, trajectories can also be reduced based on speed and direction. Changes in speed and direction can also be used to make a decision of whether a particule trajectory point should be included in a reduced trajectory or not. For our case, threshold-guided sampling approach is used that has been proposed by Potamias et. al. Safe zone is created based on the speed threshold, that is an allowable level of change in the speed of a worker. If the incoming location point resides within safe zone that is created by computing previous trajectory point`s speed and a speed threshold under consideration.Then it means, this particular point will not be contributing much information in a trajectory and hence can be discarded without compromising the accuracy of information in trajectory data. Same approach has been used to form a safe zone for a change in a direction scenario as shown in figures below. By increasing the speed or direction angle, more points will be deleted from a trajectory data.

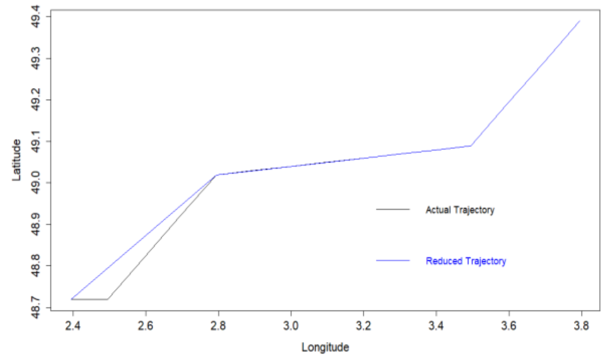

Figure 11. Reduction based on a speed 0-2.5 meter per second

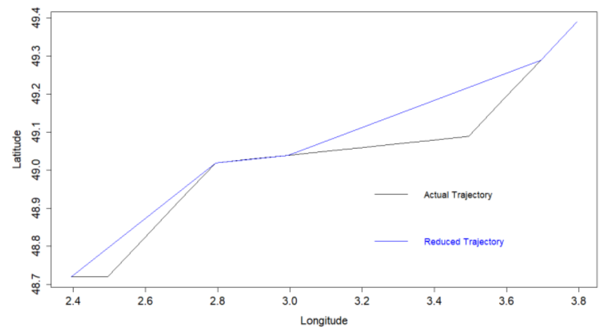

Figure 12. Reduction based on a speed 0-3.5 meter per

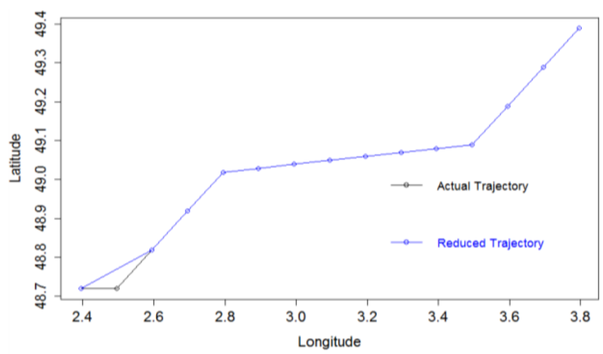

Figure 13. Reduction based on a direction angle 0-10°

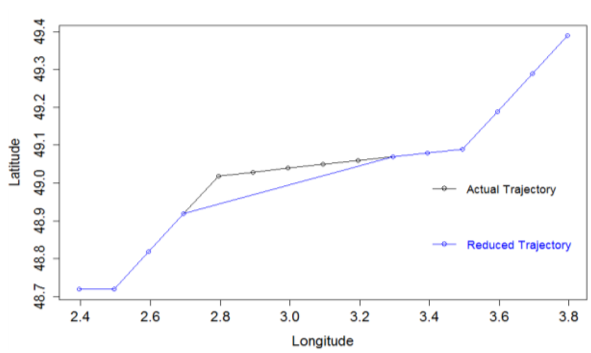

Figure 14. Reduction based on a direction angle 10-20°

5. Discussion

Secion 4 describes a system application with different data processing algorithms using positioning data of workers for safety management. If collected accurately and in real-time, the location, speed, and trajectory of workers and machineries can lead to important information regarding travel patterns. This information can then be shared among project stakeholders to improve work practices, e.g. the preconditions for safe construction operations. Additionally, the overall analyses of travel patterns on construction sites can help in the process of allocating resources more effectively to overall increase productivity and safety. Tracking and analyzing the paths of construction resources has multiple applications in construction and is not limited to important tasks such as monitoring the idle time of machines to reduce inefficiencies, tracking the path and location of materials to reduce search times, or signalling hazardous zones to workers to improve work zone safety.

There are some limitations of a proposed system application, are mentioned below;

1) The GPS data used for the system application is not a real-time data. Formerly collected CSV files of location data is used to describe preprocessing techniques of a trajectories aims to propose a new application to track worker movements on a construction site.

2) GPS technology is discussed to acquire workers` location data from a construction site. However, this technology is recommended only in the outdoor scenarios. Whereas, construction sites are composed of outdoor as well as indoor environment. Using GPS in indoor environment will introduce large measurement errors in trajectories because of interference of reflected signals. This impact will result in partial to total loss of signal tracking to capture workers` mobility. So, it is important to consider to use alternative technology to collect indoor location data.

3) The scope of the paper is kept limited to only preprocessing the geospatial trajectories. Visualizations generated using processed trajectories are still incomplete to give a complete picture of a mobility patterns of workers to building supervisors and H&S managers for decision making processed as contextual information and application domain knowledge is not been incorported.

6. Conclusion

In this paper, we have presented the importance of preprocessing spatio-temporal trajectories captured from construction sites. Once trajectory data is processed, it can be used in variety of applications for construction safety management scenarios. There is a significant gap found in the literature that application of a processed spatio-temporal trajectory data is still missing in workers`s safety monitoring scenarios in the area of construction management. To address this research gap, our paper presents a system application to track workers` movemenets on a construction site. GPS data acquired from workers` smart phones has been used and it is then preprocessed using various algorithms aims to provide effective visualizations to building supervisors and H&S manager. However, our proposed application is in the development stage lacking contextual and semantic information in the visualizations generated. Considering the preprocessing the very first step to prepare the spatio-temporal trajectories for semantic enrichment, making this processed data as a base that can leads to improved understanding of mobility of workers on construction sites. Using appropriate data and pattern mining algorithms and techniques on processed trajectory data will also help to discover interesting patterns and rules and to extract unknown and nontrival behaviors for decision-making processes.

7. REFERENCES

[1] Brillinger, D. R. (2010). Modeling spatial trajectories. Handbook of spatial statistics, 463-476.

[2] Neirotti, P., De Marco, A., Cagliano, A.C., Mangano, G. and Scorrano, F., 2014. Current trends in Smart City initiatives: Some stylised facts. Cities, 38, pp.25-36.Vancouver

[3] Jin, J., Gubbi, J., Marusic, S. and Palaniswami, M., 2014. An information framework for creating a smart city through internet of things. IEEE Internet of Things Journal, 1(2), pp.112-121.

[4] Deak, G., Curran, K. and Condell, J., 2012. A survey of active and passive indoor localisation systems. Computer Communications, 35(16), pp.1939-1954.

[5] Patroumpas, K. and Sellis, T.K., 2004, August. Managing Trajectories of Moving Objects as Data Streams. In STDBM (pp. 41-48).

[6] Andrienko, G., Andrienko, N., Bak, P., Keim, D. and Wrobel, S., 2013. Visual analytics of movement. Springer Science & Business Media.

[7] Miller, H.J. and Han, J. eds., 2009. Geographic data mining and knowledge discovery. CRC Press.

[8] Chen, M., Mao, S. and Liu, Y., 2014. Big data: A survey. Mobile Networks and Applications, 19(2), pp.171-209.

[9] Trajcevski, G., Wolfson, O., Cao, H., Lin, H., Zhang, F. and Rishe, N., 2002, April. Managing Uncertain Trajectories of Moving Objects with Domino. In ICEIS (pp. 218-225).

[10] Botea, V., Mallett, D., Nascimento, M.A. and Sander, J., 2008. PIST: an efficient and practical indexing technique for historical spatio-temporal point data. GeoInformatica, 12(2), pp.143-168.

[11] Alvares, L.O., Bogorny, V., Kuijpers, B., Moelans, B., Fern, J.A., Macedo, E.D. and Palma, A.T., 2007. Towards semantic trajectory knowledge discovery. Data Mining and Knowledge Discovery, 12.

[12] Ma, X., McCormack, E. and Wang, Y., 2011. Processing commercial global positioning system data to develop a web-based truck performance measures program. Transportation Research Record: Journal of the Transportation Research Board, (2246), pp.92-100.

[13] Šaltenis, S., Jensen, C.S., Leutenegger, S.T. and Lopez, M.A., 2000. Indexing the positions of continuously moving objects (Vol. 29, No. 2, pp. 331-342). ACM.

[14] Patroumpas, K., Alevizos, E., Artikis, A., Vodas, M., Pelekis, N. and Theodoridis, Y., 2017. Online event recognition from moving vessel trajectories. GeoInformatica, 21(2), pp.389-427.

[15] Krisnadhi, A.A., Hitzler, P. and Janowicz, K., 2016, October. A spatiotemporal extent pattern based on semantic trajectories. In Workshop on Ontology and Semantic Web Patterns, 7th edn.–WOP.

[16] Nardini, F.M., Orlando, S., Perego, R., Raffaetà, A., Renso, C. and Silvestri, C., 2018. Analysing Trajectories of Mobile Users: From Data Warehouses to Recommender Systems. In A Comprehensive Guide Through the Italian Database Research Over the Last 25 Years (pp. 407-421). Springer International Publishing.

[17] Zheng, W., Huang, X. and Li, Y., 2017. Understanding the tourist mobility using GPS: Where is the next place?. Tourism Management, 59, pp.267-280.

[18] Adey, P., 2017. Mobility. Taylor & Francis.

[19] Shen, J. and Cheng, T., 2017. Semantic Enrichment of Interesting Regions with POI data.

[20] Yan, Z., Chakraborty, D., Parent, C., Spaccapietra, S. and Aberer, K., 2013. Semantic trajectories: Mobility data computation and annotation. ACM Transactions on Intelligent Systems and Technology (TIST), 4(3), p.49.

[21] Mousavi, A., 2016. Ontology Based Semantic Knowledge Discovery for Movement Behaviors (Doctoral dissertation, University of Calgary).

[22] Ali, S., Khusro, S., Ullah, I., Khan, A. and Khan, I., 2017. SmartOntoSensor: Ontology for Semantic Interpretation of Smartphone Sensors Data for Context-Aware Applications. Journal of Sensors, 2017.

[23] Chakri, S., Raghay, S., & El Hadaj, S. (2017). A Semantic Interpretation of Unusual Behaviors Extracted from Outliers of Moving Objects Trajectories. INTERNATIONAL JOURNAL OF ADVANCED COMPUTER SCIENCE AND APPLICATIONS, 8(3), 119-127.

[24] MARKETOS, G., FRENTZOS, E., NTOUTSI, I., PELEKIS, N., RAFFAETA, A., AND THEODORIDIS, Y. 2008. Building real world trajectory warehouses. In Proceedings of the 7th ACM International Workshop on Data Engineering for Wireless and Mobile Access (MobiDE'08). 8–15.

[25] JUN, J., GUENSLER, R., AND OGLE, J. 2006. Smoothing methods to minimize impact of global positioning system random error on travel distance, speed, and acceleration profile estimates. Transport. Res. Rec. J. Transport. Res. Board 1972, 1, 141–150.

[26] Richter, K. F., Schmid, F., & Laube, P. (2012). Semantic trajectory compression: Representing urban movement in a nutshell. Journal of Spatial Information Science, 2012(4), 3-30.

[27] Schmid, F., Richter, K. F., & Laube, P. (2009). Semantic trajectory compression. Advances in Spatial and Temporal Databases, 411-416.

[28] MERATNIA, N. AND DEBY, R. A. 2004. Spatiotemporal compression techniques for moving point objects. In Proceedings of the 9th International Conference on Extending Database Technology (EDBT'04). 765–782.

[29] BUCHIN, M., DRIEMEL, A., KREVELD, M. V., AND SACRISTAN, V. 2010. An algorithmic framework for segmenting trajectories based on spatio-temporal criteria. In Proceedings of the 18th Annual ACM International Symposium on Advances in Geographic Information Systems (GIS'10). 202–211.

[30] Buchin, M., Driemel, A., Van Kreveld, M. and Sacristán, V., 2011. Segmenting trajectories: A framework and algorithms using spatiotemporal criteria. Journal of Spatial Information Science, 2011(3), pp.33-63.

[31] Hu, Yingjie, Krzysztof Janowicz, David Carral, Simon Scheider, Werner Kuhn, Gary Berg-Cross, Pascal Hitzler, Mike Dean, and Dave Kolas. "A geo-ontology design pattern for semantic trajectories." In International Conference on Spatial Information Theory, pp. 438-456. Springer, Cham, 2013.

[32] Yan, Z., Chakraborty, D., Parent, C., Spaccapietra, S., & Aberer, K. (2011, March). SeMiTri: a framework for semantic annotation of heterogeneous trajectories. In Proceedings of the 14th international conference on extending database technology (pp. 259-270). ACM.

[33] Wagner, R., de Macedo, J. A. F., Raffaetà, A., Renso, C., Roncato, A., & Trasarti, R. (2013, November). Mob-warehouse: A semantic approach for mobility analysis with a trajectory data warehouse. In International Conference on Conceptual Modeling (pp. 127-136). Springer, Cham.

[34] Noël, D., Villanova-Oliver, M., Gensel, J., & Le Quéau, P. (2015, November). Modeling semantic trajectories including multiple viewpoints and explanatory factors: application to life trajectories. In Proceedings of the 1st International ACM SIGSPATIAL Workshop on Smart Cities and Urban Analytics (pp. 107-113). ACM.

[35] Vandecasteele, A., Devillers, R., & Napoli, A. (2014). From movement data to objects behavior using semantic trajectory and semantic events. Marine Geodesy, 37(2), 126-144.

[36] Brisaboa, N. R., Luaces, M. R., Pérez, C. M., & Places, Á. S. (2017, May). Semantic Trajectories in Mobile Workforce Management Applications. In International Symposium on Web and Wireless Geographical Information Systems (pp. 100-115). Springer, Cham.

[37] Wagner, R., de Macedo, J. A. F., Raffaetà, A., Renso, C., Roncato, A., & Trasarti, R. (2013, November). Mob-warehouse: A semantic approach for mobility analysis with a trajectory data warehouse. In International Conference on Conceptual Modeling (pp. 127-136). Springer, Cham.

[38] Li, Y., Su, H., Demiryurek, U., Zheng, B., He, T., & Shahabi, C. (2017, April). PaRE: A System for Personalized Route Guidance. In Proceedings of the 26th International Conference on World Wide Web (pp. 637-646). International World Wide Web Conferences Steering Committee.

[39] Zhang, J. (2014). Efficient frequent sequence mining on taxi trip records using road network shortcuts. Big Data: Techniques and Technologies in Geoinformatics, HA Karimi, Ed, 193-206.

[40] Chen, C., Su, H., Huang, Q., Zhang, L., & Guibas, L. (2013, November). Pathlet learning for compressing and planning trajectories. In Proceedings of the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (pp. 392-395). ACM.

[41] Bogorny, V., Renso, C., Aquino, A. R., Lucca Siqueira, F., & Alvares, L. O. (2014). Constant–a conceptual data model for semantic trajectories of moving objects. Transactions in GIS, 18(1), 66-88.

[42] Lv, M., Chen, L., & Chen, G. (2012, October). Discovering personally semantic places from GPS trajectories. In Proceedings of the 21st ACM international conference on Information and knowledge management (pp. 1552-1556). ACM.

[43] Zheng, Y., Zhang, L., Xie, X. and Ma, W.Y., 2009, April. Mining interesting locations and travel sequences from GPS trajectories. In Proceedings of the 18th international conference on World wide web (pp. 791-800). ACM.

[44] Y. Zheng, Q. Li, Y. Chen, and X. Xie. 2008a. Understanding mobility based on GPS data. In Proceedings of the 11th International Conference on Ubiquitous Computing. ACM, 312–321.

[45] Y. Zheng, L. Liu, L. Wang, and X. Xie. 2008b. Learning transportation mode from raw GPS data for geographic application on the Web. In Proceedings of the 17th International Conference on World Wide Web. ACM, 247–256.

[46] Sun, X., Wandelt, S., Hansen, M. and Li, A., 2017. Multiple airport regions based on inter-airport temporal distances. Transportation Research Part E: Logistics and Transportation Review, 101, pp.84-98.

[47] Michalis Potamias, Kostas Patroumpas, and Timos Sellis. Sampling Trajectory Streams with Spatiotemporal Criteria. In SSDBM ’06: Proceedings of the 18th International Conference on Scientific and Statistical Database Management, pages 275–284. Ieee, 2006. 23, 67, 134.