Understanding Worker Mobility within the Stay Locations using HMMs on Semantic Trajectories

Muhammad Arslan
Univ. Bourgogne Franche-Comte,
Batiment I3M rue Sully, Dijon, France
muhammad.arslan@u-bourgogne.fr

Christophe Cruz
Univ. Bourgogne Franche-Comte,
Batiment I3M rue Sully, Dijon, France
christophe.cruz@ubfc.fr

Dominique Ginhac
Univ. Bourgogne Franche-Comte,
Batiment I3M rue Sully, Dijon, France
dominique.ginhac@ubfc.fr

Abstract—**The analysis of dynamic interactions of objects has always received great attention due to widespread applications of positioning technologies such as Indoor Positioning Systems (IPS) or the Global Positioning System (GPS). The outcomes of such analysis lead towards a better understanding of workers` behaviors for mobility recognition that can enhance safety in construction operations. For capturing dynamic interactions of workers on a construction site, a solution based on semantic trajectories and Hidden Markov Model (HMM) is presented in the form of three subsystems. First, the pre-processing subsystem constructs trajectories using a stream of raw spatio-temporal points captured from an IPS system. Then, stay locations of workers are identified along with their frequency of visiting to understand the importance of such locations in daily use. In the second subsystem, an ontology-based STRiDE (Semantic Trajectories in Dynamic Environments) model is applied that provides a data structure for transforming raw trajectories into semantic trajectories. The model performs semantic enrichment processes by using semantic data sources including application domain knowledge and geographic databases to allow the understanding of meanings behind workers` mobility behaviors. Lastly, the third subsystem executes the HMM along with the Viterbi algorithm to categorize different mobility behaviors of workers within the identified stay locations. Based on these subsystems, visualizations of semantic trajectories and mobility patterns are generated to help facility managers in monitoring and controlling site activities by identifying high-risk workers` behaviors.**

Keywords—semantic trajectories; stay locations; GPS; mobility; Hidden Markov Models; worker safety

I. INTRODUCTION

Harsh conditions, frequent involvement of high safety risks and constantly changing environment make construction sites hazardous for workers [1]. If compared with other industries, construction industry has one of the highest accidents rates [2]. According to the Bureau of Labor Statistics (BLS), in 2016 out of 5,190, fatal work injuries, 19 % of fatalities were recorded from the U.S. construction industry [2]. A closer look into BLS statistics indicates that unsafe human behaviors, difficult site conditions and workers striking against or being struck by moving objects were the leading causes of the fatal accidents [3]. These causes highlight the need for more efficient construction resources` mobility monitoring systems for their unsafe behaviors.

Time series trajectory data [4] acquired from smartphones equipped with positioning technologies such as GPS/IPS are crucial to understanding the dynamics of moving objects on construction sites. Thus, knowledge extracted from such data can help to find existing relevant patterns [4] to model the complex and unpredictable dynamic building environment. A raw time series trajectory data is a collection of an ordered sequence of spatio-temporal triples in the form of ⟨latitude, longitude, timestamp⟩ extracted at equally spaced time intervals from the location data of an object moving in a geographical space [4, 5]. Raw trajectories acquired from location acquisition devices are well-suited for applications which aim to find the location of a moving object or to calculate statistics such as distance, direction, acceleration, etc. [9]. However, additional information from the application context can be used to make raw trajectories more meaningful for supporting advanced trajectory applications. The additional pieces of information are called as annotations which can be collected using different external data sources such as web pages, applications, and geo-databases [5]. The process of adding annotations to a trajectory as a whole or its subpart is known as semantic enrichment [4, 5]. Once the trajectory data is semantically enriched, it offers remarkable opportunities to predict workers` movements by understanding their mobility patterns. However, to make efficient use of the trajectory data in a comprehensive way, several pre-processing techniques [5] should be applied first to transform raw trajectories into semantically processed trajectories. After semantic enrichment, HMMs along with the Viterbi algorithm can be applied to extract trajectory patterns [6]. HMMs are statistical tools to represent probability distributions (PDs) over sequences of observations and widely used for predictive analytics [6]. The extracted patterns can be useful for the facility managers to visualize mobility of workers which can lead to improved facility management intervention strategies and also help in identifying high-risk workers behaviors to prevent safety hazards.

The paper is organized as follows. Section 2 introduces the background literature used for this work. Section 3 presents the proposed solution to extract mobility patterns from semantic trajectories using HMMs. Section 4 describes the limitations of the presented work, a conclusion and some future works.

II. BACKGROUND

The literature review was initially performed to methodically collect information for identifying and understanding the problem domain, which is to extract mobility patterns of construction workers from their semantically enriched trajectories after finding their stay locations. At first, a review of the literature on smart buildings reveals that the trajectory data for understanding fine-grained user-building interactions is only useful if the data model has an ability to store information of moving and changing building objects [7]. Trajectory data

modeling is the first step to design as it should have capability to provide not only the spatio-temporal view but also a semantic view by explaining the underlying meanings behind trajectories [4, 5].

A. Trajectory Modeling

It is a process to define data requirements to support trajectories' applications [8, 9]. Frihida et al. [10] presented a model to represent trajectories as an Abstract Data Type (ADT). The model integrates spatial, temporal, and thematic dimensions for representing and manipulating the trajectory data. The major limitation of abstract data type-based approach is the dependency of the trajectory data type on the application as model represents trajectories as a series of connected trips and activities. Data type-based models alone are not enough for constructing trajectory applications because it imposes the use of a generic data type to represent trajectories for all the applications [11]. To address this drawback, Parent et al. designed pattern-based model [12], that is based on Model Analysis and Decision Support (MADS). It supports spatial and temporal objects and relationships by providing a method for describing the spatial extent and lifespan of the trajectory. However, to define trajectory components in 'stop' and 'move' segments, there is a need to manually input the contextual information in the model according to the application [11].

Ontology is the conceptualization of a specific domain for showing relationships between concepts in the form of a hierarchy [11]. An example of a model based on ontologies is presented by Yan and Spaccapietra [13]. It offers a multilayered model to represent trajectories. In addition, Noël et al. [14] presented ontological design patterns to model trajectories for understanding trajectory events in a better way. If ontology-based modeling is compared with formally discussed two modeling approaches, it represents richer semantic information by integrating different types of information enrichment processes [11]. Existing literature also presents solutions based on hybrid modeling approach to combine best features of different models discussed above for constructing trajectories [11]. The hybrid model presented by Yan et al. [4] offers different levels of data abstraction by encapsulating geometry and semantics of trajectories. The important aspect of using the hybrid approach is to cover requirements of a variety of applications by offering three different types of data abstraction that are processing raw trajectories, constructing trajectory episodes, and generating semantically enriched trajectories.

B. Trajectory Preprocessing

Once location data of moving objects is acquired and data requirements for modeling trajectories are fixed, raw trajectories should be pre-processed according to the application requirements [15]. The basic idea of pre-processing trajectories is to reduce the storage, processing, and communication overheads without compromising the precision of a trajectory data [5]. The basic tasks of pre-processing include noise filtering, reduction, segmentation, and stay point detection [5, 15]. The objective of noise filtering is to remove noise from trajectories caused by weak signals of location acquisition systems. Trajectory reduction processes are used to reduce the size of the trajectory to minimize the computation overhead while keeping the usefulness of the trajectory [5]. Segmentation

deals with the algorithms to divide trajectories into different segments. The criteria of segmentation can be based on the time interval, spatial properties or semantic meaning [5]. Stay point detection techniques identify the location points where the moving object has spent some time by staying over there within a specified distance. A stay point location can be inside the building or outside such as a shopping mall, a restaurant or an office.

C. Trajectory Semantic Enrichment

After pre-processing of trajectories, the process of annotation is applied to transform IPS/GPS trajectories into semantic trajectories. In this process, trajectory segments having specific intervals are assigned meaningful information regarding their mobility tracks [9]. As trajectories are already divided into stop and move episodes during the segmentation process, now it's important to annotate stops. Annotation is a process to add application context data to raw spatio-temporal trajectories [5, 9]. A conventional semantic enrichment process for trajectories receives a collection of raw trajectories as an input and produces annotated trajectories, which are known as semantic trajectories [9]. The annotation process involves mapping episodes to the meaningful information such as; the mapping of places of interest that can be in the form of Points of Interest (POI), the region of interest (ROI), roads in the form of lines or the geographical regions as discussed in [9].

The STriDE [7] based on a Continuum model [16 - 21] is a type of ontology-based trajectory data modeling [13, 14] is applied for performing semantic enrichment process [9]. After recognizing the key regions in the building as stay locations in the form of POIs, behavior patterns are extracted for understanding mobility within these critical locations. Probabilistic HMM along with the Viterbi algorithm as discussed below is applied to categorize the mobility into safe and unsafe patterns.

D. Hidden Markov Models (HMMs)

HMM is a statistical tool to represent probability distributions (PDs) over sequences of observations and widely used for predictive analytics [22]. Generally, the HMM has three main properties that defines it [6]. The first, it assumes that the observation was generated by a process at time t whose state S_t is hidden from the observer. The second, it assumes that the state of this hidden process fulfills the Markov property. That is, the current state S_t of the process is only dependent on the only previous state S_{t-1} and independent of all the states prior to $t-1$. It means that to predict the future of the process, the hidden state encapsulates all the information we need to know from the process history. The third supposition of the HMM is that the hidden state variable can only take T integer values $\{1, 2, 3, ..., T\}$. In general, an HMM λ is described by a set of three parameters which can be written as the 3-tuple $\lambda = (A, B, \pi)$. Here, A is the transition probability matrix, B is the emission probability matrix, and π is the vector of the initial state probabilities. There are three basic scenarios for which HMM can be used for real-world applications [6]. These scenarios are the following; 1) Computing the probability of the observation sequence $P(O|\lambda)$ using the given HMM (λ) and the observation sequence $O = \{O_1, O_2, O_3, ..., O_T\}$. 2) Extracting the

most optimal hidden state sequence ($Q = q_1, q_2, q_3 \dots q_T$) which best explains the observations using the given HMM (λ) and the observation sequence $O = \{O_1, O_2, O_3, \dots, O_T\}$. 3) Adjusting the values of the state transition probabilities (A) and the output emission probabilities (B) of the HMM (λ) to maximize the probability of the observation sequence $O = \{O_1, O_2, O_3, \dots, O_T\}$. This process is called training as the observation sequence is used to train the HMM to make it best for observing real phenomena.

However, this work focuses on the latter two scenarios.

III. Semantic Trajectory Analytics using HMMs

From our knowledge, no previous work was done on understanding workers` mobility patterns within the stay locations using HMMs on semantic trajectories. Thus, we have developed a prototype system which uses a data model STriDE [7] to store trajectories after applying pre-processing techniques. For proof-of-concept application, trajectory data of users that are working inside the already constructed building is acquired in real-time, then semantically enriched and their mobility patterns are extracted within their stay locations for experimental analysis as mentioned below. However, the functionality of the system will remain the same if deployed on a construction site.

A. Trajectory Pre-processing and Semantic Enrichment

To understand the interactions of users in the dynamic environment, use of IPS trajectory data is considered. For the trajectory data acquisition, 200 Bluetooth beacons by Kontakt IO Company (https://kontakt.io) are placed in different locations in the building. These beacons work using the triangulation technique to estimate user positions. The longitude and latitude coordinates are generated after the pre-processing done by the mobile application. In a day, 13,223 locations points have been recorded across different locations with the sampling interval of 1 minute. A single IPS data log consists of a user identification (ID), timestamp, elevation, latitude, and longitude value. After collecting location data, it will undergo the preprocessing steps for transforming into trajectories.

The first step is to use an appropriate filter to remove outliers from the data. There exist many outliers removing filters to improve the data quality [5]. However, the median filter is used on an acquired IPS data because it depicts robustness property in filtering (as shown in Fig. 1a) [5]. After removing the outliers, stay points of workers are calculated to enrich a trajectory with semantic points in the form of stop and move segments using the Zheng et al. approach [5]. The main purpose of calculating stay points in trajectories is to find locations in the building where workers are spending most of their time. This information will help to track the occurrence of unexpected situations in the building if the stay duration is greater or less than the required. By setting the distance (D_{thresh}) value to 3 meters and time threshold (T_{thresh}) value to 20 minutes, stay points in a trajectory are identified as shown in Fig. 1b. However, these thresholds are totally application dependent and can be changed according to the indoor or outdoor environment. In addition, the frequency of visiting the same stay locations is also calculated to understand which stay locations are visited more often by the workers.
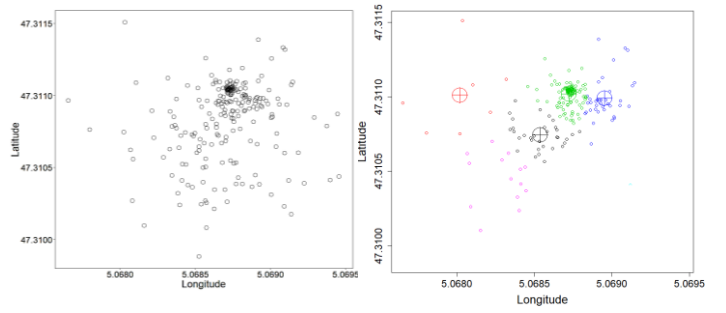

Fig. 1 a) Trajectory data of a user. b) Detection of stay points in different colors

Once stay points are identified, there is a need to label each stay point with a building identification (ID) that corresponds to different building locations such as Room106. For tagging building identification (ID), we have used STriDE model for enriching trajectory data with relevant semantic information that corresponds to the building architecture.

STriDE model requires an OpenStreetMap (OSM) file for a building under analysis, a set of semantic rules, and a taxonomy (e.g. a BIM file [21]). OSM provides spatially rich geographic vector data in an XML format that is open and well-documented having the authorization for utilizing, copying or editing it. OSM data model contains nodes, ways, and relations. In addition, an OSM entity can be tagged with single or multiple key-value pairs to add information. This information can be as simple (e.g. label) or complex (derived information from safety instructions) as per the application`s requirements. These key-value pairs are clustered and maintained with the help of taxonomies created by the domain experts. In our case, an OSM file is used in which boundaries of the rooms are defined along with their links to each other. The file carries the complete building plan and its surroundings for defining the building structure. For labeling, each building location, a taxonomy is created in accordance with the building requirements. The created taxonomy is a hierarchy of concepts written as RDF triples using SKOS (Simple Knowledge Organization System) vocabulary. In addition, semantic rules are constructed in the form of a JSON file to link each OSM object with the taxonomy. Then, OSM file along with semantic rules are fed to a 2-step Java parser. Firstly, the parser will establish the mapping between each OSM entity with the Java object. Secondly, using semantic rules, these constructed objects are transformed into new Java objects according to the semantic definition (see Fig. 2). These processed objects are later stored in a Stardog triple store (www.stardog.com) to achieve the complete representation of the building environment. In the Fig. 2, geometry is defined outside the main entity. Here, entity is describing the identity of the building location. For our scenario, on construction sites, the geometry of locations is changing over time. The purpose of using the STriDE model is that, it allows us to log changes in shape, size, and attributes of dynamic entities using the concept of timeslices. A timeslice consists of four components: an identity, properties having alphanumeric values, a time component indicating the validity of the timeslice, and a geographical component depicting the spatial representation of the entity. In case of any change in the components of the time slice, excluding the identity, a new time slice is created inheriting the components of the last known state of the entity and then those changes are applied. This newly created time slice is then linked

with the previous one to keep the evolution of the entity during its life cycle.

```
stride:W235 a stride:Entity ;
    skos:prefLabel "Corridor of floor 0" ;
    rdfs:comment "" .

stride:GEO-W235-0 a geo:Geometry ;
    geo:asWKT "LINESTRING (20 65, 15 65, 15 50, 20 50, 15 50, 15 15, 20
10, 15 15, 30 15, 30 20, 30 15, 50 15, 50 20)"^^geo:wktLiteral .

stride:W235-0 a stride:TimeSlice ;
    stride:hasStartDate "2018-01-01T00:00:00"^^xsd:dateTime ;
    stride:hasFeature stride:W235 ;
    geo:hasGeometry stride:GEO-W235-0 ;
    dct:subject stride:Corridor ;
    stride:hasEndDate "9999-12-31T23:59:59"^^xsd:dateTime .
```

Fig. 2. Parsed OSM file using the semantic rules and the taxonomy (The script is the RDF Turtle definition of an object of the kind "Corridor" identified by the value stride:W235. These values are used to define semantic trajectories.)

After mapping the locations in a trajectory to building IDs, a trajectory is enriched with semantic lines in the form of movement types which are being carried by the users by calculating the speed values between two points in meter/second. Finally, mapping is done with the semantic region (building architecture) as shown in Fig. 3. Using the STriDE model, our system stores spatio-temporal trajectories of moving objects after data cleaning, segmentation, and semantic enrichment processes. These semantically enriched processed trajectories are later used for the extraction of mobility-related behavior information by applying the HMM probabilistic framework as discussed in the next section.
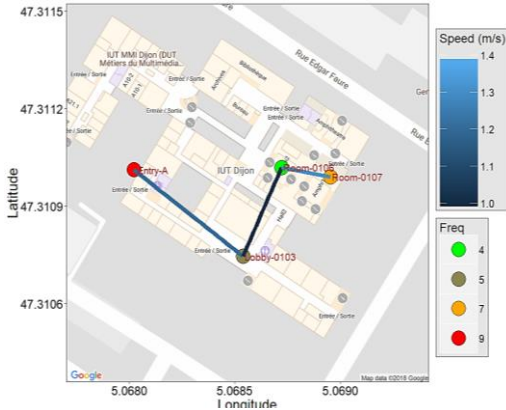

Fig. 3. The enriched user trajectory is composed of stay points in three different colors (frequency of visiting (no. of times in a time threshold)), semantic lines connecting stay points (based on speed values calculation), and a semantic region labeling stay points (building structures, room identification for e.g. Room0106).

B. Extracting Mobility Patterns using HHMs

To categorize the mobility data of workers into different states for each stay location, an independent HMM is trained for patterns extraction. According to the existing literature [23, 24], the mobility of a moving object can easily be defined using individual steps and turning angles. Before feeding the trajectory sequences into the HMM, we need to visualize the data in order to set the hidden state values using measurement variables. For our case, we have used values of 'step length' (distance between two trajectory points) and 'turning angle' (change in direction in radians from the previous point to the current point), which were

extracted from pre-processing the trajectories. Step length (l_t) is calculated using the Haversine distance formula [25] between the locations (x_t, y_t) and (x_{t+1}, y_{t+1}) as below;

$$d = 2r sin^{-1} \left(\sqrt{sin^2 \frac{x_t - y_t}{2} + cos\, x_t\, cos\, x_{t+1}\, sin^2 \frac{y_t - y_{t+1}}{2}} \right.$$

Where "r" is the radius of the earth. The reason for using the Haversine distance formula is because it is the one of the preferred method that calculates the geographic distance between two points on a sphere [26]. While turning angle (ϕ_t) is calculated as the change in bearing (b_t) as $(b_t) = atan2(y_{t+1} - y_t, x_{t+1} - x_t)$ between the time intervals $[t - 1, t]$ and $[t, t + 1]$. This step will help us to detect any outliers left in the trajectory.

For mobility categorization of the workers within the stay locations, we need to select the hidden states. This is the most important step while training the HMM to generate patterns. From the existing literature [26], the normal walking speed for an adult range from 1.0 to 1.6 meters per second (m/s). However, keeping an indoor environment into the consideration, we have used the value of 1.4 m/s as a safe walking speed limit that will give us 84 steps per minute i.e. the sum of step lengths for a minute. Three hidden states are defined using different values of 'step length' and 'turning angle' are; 'static', 'normal', and 'risky' as shown in Table I. The number of hidden states in our case is 3. However, the number of states can be decreased or increased according to the application requirements. The purpose of defining such hidden states is to understand different types of workers` mobility within the stay location. To define hidden states in the HMM, there is a need to specify distributions according to the trajectory data. For this, we have used gamma distribution for step lengths and Von Mises distribution (also known as the circular normal distribution) for turning angles.

TABLE I CRITERIA FOR DEFINING HIDDEN STATES

State/ symbol	Step length threshold (No. of steps in a minute)	Turning angle threshold (Radians)
Static S_1	0	angle $< \pi/2$
Normal S_2	$0 <$ steps ≤ 84	$\pi/2 \leq$ angle $< \pi$
Risky S_3	steps > 84	angle $\geq \pi$

After defining the hidden states, we have used Baum-Welch algorithm [6] in the R studio that allows learning parameters of an HMM using the maximum likelihood approach and computes λ^* that maximizes the likelihood of the sample of training sequences $\chi = \{O^k\}_k^K$ namely $P(\chi|\lambda)$. After computing the transition probabilities of workers to move between different hidden states using the Baum-Welch algorithm, most probably occurring patterns are also extracted based on the output of learned parameters. For this, we have used Viterbi algorithm [6], a form of dynamic programming to extract the most probable sequence of states for a given trajectory. For each symbol o_T in the observation trajectory O, an algorithm calculates the probability of its emission for each possible hidden state. The algorithm starts by computing the initial probability of the emission of the symbol o_1 in all possible states. Then, it

computes again the emission of the symbol o_2 for each state transition. This process is repeated for every observation symbol until the observation sequence ends that is at step T. Finally, having all possible paths covered, the Viterbi algorithm [6] look for the path and output the most probable state sequence $Q = \{q_1, q_2, q_3 \ldots q_T\}$. Fig. 4 shows the most probable state patterns within a stay location using the trained HMM.

C. Discussion and Modeling Checking

The developed system identified 4 stay locations in the building. However, for the proof of concept we have used a single stay location for patterns analysis. Though, the same process can be repeated for each stay location to analyze all stay locations. As it can be seen from the extracted patterns (see Fig. 4) that there are fluctuations in the mobility behaviors observed during the different times of a day. Visualizing such patterns in real-time will help facility managers to monitor stay locations for identifying accident prone scenarios on the construction sites and enabling them to take quick actions in case of detection of 'risky' worker behaviors. It is also important to measure the general goodness of fit of the trained model that has been used for extracting patterns. The trained model should depict that it was a true data generating process of trajectory observations which have been fed into the HMMs during the training process. For this, the pseudo-residuals also known as quantile residuals of the trained HMM are computed for the model checking. The pseudo-residuals follow a standard normal distribution if the trained model is a true data-generating process. It means that, if the model fits the data well, the points in the qq-plot will be closer to the straight line and a deviation from normality will indicate a lack of fit. Moreover, the absolute value of residual increases with increasing deviation from the straight line and specific observations can be known in less time. For more details on pseudo-residuals, see Zucchini et al. [27]. The pseudo-residuals of the 3-state model fitted to the trajectory data are displayed in Fig. 5. The plot shows that the trained model is well fitted for the observation dataset and has few deviations from the normality. However, the goodness of fit of the model can be improved by increasing the number of hidden states. The dataset is available online (https://github.com/ChristopheCruz/LivingLabStride).

IV. CONCLUSION AND FUTURE WORK

The study explained the functionality of the developed prototype system that uses the IPS technology and HMM based probabilistic approach to identify unsafe worker behaviors within the stay locations. For the trajectory data acquisition, Bluetooth beacons are used. Whereas, trajectory enrichment is achieved using the STriDE model. Though, Bluetooth beacons are not recommended as an indoor tracking solution because of less precision in the accuracy of determining the locations. These errors are noticeable in Fig. 3 as collected trajectory data points did not completely join spatially with the semantic region information extracted from the OSM file. Consequently, the closest possible semantic regions have been mapped. For extracting the stay locations, the subsequent trajectory data points at the same location are grouped together. For grouping, a time threshold of 20 minutes is used. It means that subsequent stops between which the time gap is less than 20 minutes are grouped together. The reason for the 20 minutes threshold is that

intervals smaller than 20 minutes are likely to represent a worker that was just visiting the stay location for a short period of time. After stay location detection and semantic enrichment, HMM along with the Viterbi algorithm is used to estimate mobility patterns within the stay locations after training the model. The developed system will not only serve as a semantic trajectory visualization platform for facility managers but will provide most probable occurring mobility patterns that can improve building monitoring processes by identifying unsafe worker behaviors. As the proposed system prototype is still in the development phase and this article presents some initial results. Further work needs to be done to take a complete advantage of the STriDE model for modeling and testing the dynamic environment scenarios where the purpose of locations is constantly changing with the passage of time (for example, a work room is now a storage room). The change in the purpose of the room will result in different stay locations having different worker mobility patterns. These dynamic changes and interactions can be efficiently managed using the STriDE model that is designed in a way to capture the information of changing as well as moving objects in a dynamic building scenarios. Moreover, future research needs to focus to empirically test the system on real construction sites as at present the system is modelled for an already constructed building. Taking information from the real-time site cases will improve the training of HMMs. It will result in improved and precise mobility patterns extraction for capturing fast moving and turning construction resources such as workers and machineries on construction sites to reduce accidents.

REFERENCES

[1] R. A. Haslam, S. A. Hide, A. G. Gibb, D. E. Gyi, T. Pavitt, S. Atkinson, and A. R. Duff, "Contributing factors in construction accidents," Applied ergonomics, 36(4), 2005, pp. 401-415.

[2] Stats.bls.gov.2017, "Census of Fatal Occupational Injuries (CFOI) - current and revised data," available at: http://stats.bls.gov/iif/oshcfoi1.htm (Accessed 28 March 2018).

[3] M.Arslan, C. Cruz and D. Ginhac, "Semantic enrichment of spatio-temporal trajectories for worker safety on construction sites," Procedia computer science, 130, 2018, pp.271-278.

[4] Z. Yan, D. Chakraborty, C. Parent, S. Spaccapietra and K. Aberer, "Semantic trajectories: Mobility data computation and annotation," ACM Transactions on Intelligent Systems and Technology (TIST), 4(3), 2013, pp. 1-38.

[5] Y. Zheng, "Trajectory data mining: An overview," Transactions on Intelligent Systems and Technology, 6(3), 2015, pp.1-41.

[6] L. R. Rabiner, "A tutorial on Hidden Markov Models and selected applications in speech recognition," Proceedings of the IEEE, 77(2), 1989, pp. 257-286.

[7] C. Cruz, "Semantic trajectory modeling for dynamic built environments," IEEE International Conference on Data Science and Advanced Analytics (DSAA), 2017, pp. 468-476.
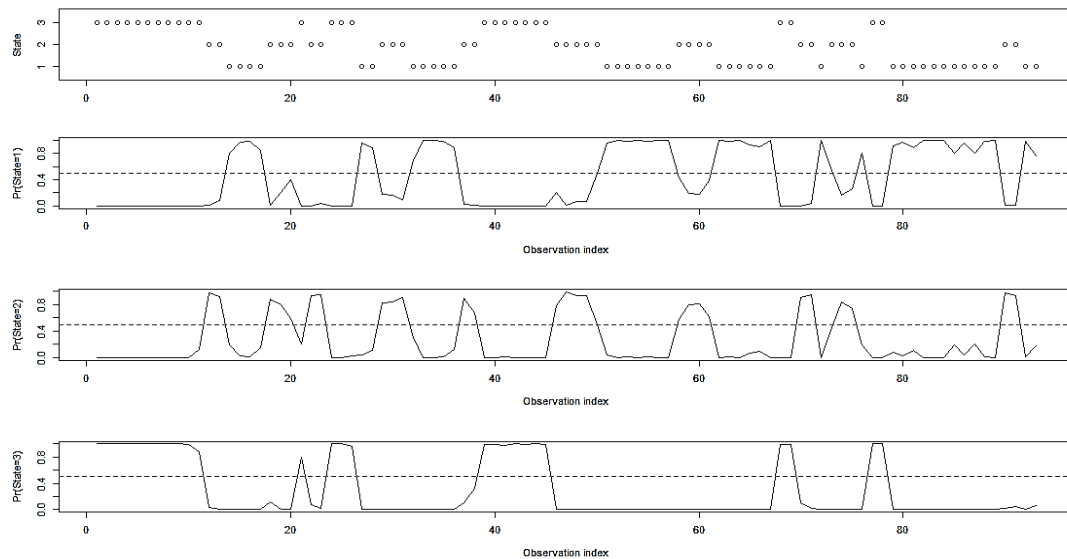
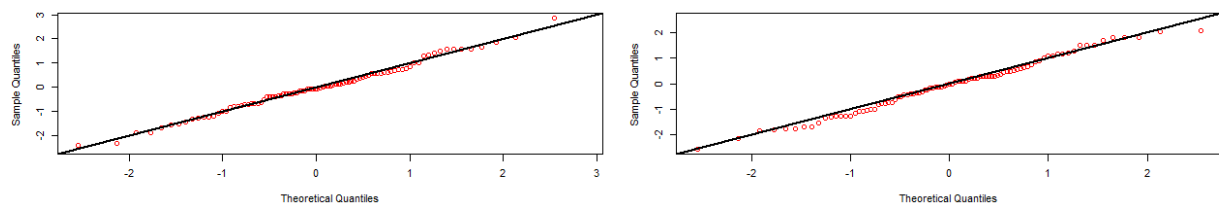Fig. 4 Decoded states sequence (top row), and state probabilities of observations (last three rows)

Fig. 5 qq-plots of the pseudo-residuals of the 3-state HMM

[8]   S. Spaccapietra, C. Parent, M.L. Damiani, J.A. de Macedo, F. Porto and C. Vangenot, "A conceptual view on trajectories," Data and Knowledge Engineering, 65(1), 2008, pp.126-146.

[9]   F. Wu, H. Wang, Z. Li, W.C. Lee and Z. Huang, "SemMobi: A semantic annotation system for mobility data," Proceedings of the 24th International Conference on World Wide Web (WWW '15 Companion), 2015, pp. 255-258.

[10]   A. Frihida, D. Zheni, H.B. Ghezala and C. Claramunt, "Modeling trajectories: A spatio-temporal data type approach," 20th International Workshop on Database and Expert Systems Application, DEXA'09, 2009, pp. 447-451.

[11]   B.H. Albanna, I.F. Moawad, S.M. Moussa and M.A. Sakr, "Semantic trajectories: a survey from modeling to application," Information Fusion and Geographic Information Systems (IF&GIS'2015), Springer International Publishing, 2015, pp. 59-76.

[12]   C. Parent, S. Spaccapietra and E. Zimányi, Conceptual modeling for traditional and spatiotemporal applications: the MADS approach, Springer, New York, 2006.

[13]   Z. Yan and S. Spaccapietra, "Towards semantic trajectory data analysis: a conceptual and computational approach," VLDB PhD Work 15(2), 2009.

[14]   D. Noël, M. Villanova-Oliver, J. Gensel and P. Le Quéau, "Design patterns for modelling life trajectories in the semantic web," International Symposium on Web and Wireless Geographical Information Systems, 2017, pp. 51-65.

[15]   C. Parent, S. Spaccapietra, C. Renso, G. Andrienko, N. Andrienko, V. Bogorny, M.L Damiani, A. Gkoulalas-Divanis, J. Macedo, N.Pelekis, and Y.Theodoridis, "Semantic trajectories modeling and analysis," ACM Computing Surveys (CSUR), 45(4), 2013, pp.1-37.

[16]   B.Harbelot, H. Arenas and C. Cruz, "LC3: A spatio-temporal and semantic model for knowledge discovery from geospatial datasets," Web Semantics: Science, Services and Agents on the World Wide Web, 35(1), 2015, pp. 3-24.

[17]   B. Harbelot, H. Arenas and C. Cruz, "Continuum: A spatiotemporal data model to represent and qualify filiation relationships," Proceedings of the 4th ACM SIGSPATIAL International Workshop on GeoStreaming (IWGS), 2013.

[18]   H. Arenas, B. Harbelot and C. Cruz, "A Semantic analysis of moving objects using as a case study maritime voyages from eighteenth and nineteenth centuries," The 6th International Conference on Advanced Geographic Information Systems, Applications, and Services, 2014.

[19]   B. Harbelot, H. Arenas and C. Cruz, Using semantic web technologies to follow the evolution of entities in time and space," International J. On Advances in Intelligent Systems, Freimut Bodendorf, University of Erlangen-Nuernberg, Germany, 2013, pp.256-265.

[20]   B. Harbelot, H. Arenas, Helbert and C. Cruz, LC3: A spatio-temporal and semantic model for knowledge discovery from geospatial datasets, Web Semantics: Science, Services and Agents on the World Wide Web, vol. 35, 2015, pp 3-24.

[21]   R. Vanlande, C. Nicolle, C. Cruz, IFC and Buildings Lifecycle Management, J. of Automation in Construction, Elsevier, 2008

[22]   C. M. Grinstead and J. L. Snell, Introduction to probability, American Mathematical Society, 2012.

[23]   Ľ.Ilkovičová, P. Kajánek and A. Kopáčik, "Pedestrian indoor positioning and tracking using smartphone sensors step detection and map matching algorithm," International Symposium on Engineering Geodesy, 2016, pp. 20-22.

[24]   C. Andolfo and F. Sadeghpour, "A probabilistic accident prediction model for construction sites," Procedia Engineering, 123, 2015, 15-23.

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

[26]   S. Harkema, A. Behrman and H. Barbeau, Chapter 16 - Evidence-based therapy for recovery of function after spinal cord injury, Handbook of Clinical Neurology, Elsevier, 109, 2012.

[27]   W. Zucchini, I. L. MacDonald and L. Langrock, Hidden Markov Models for time series: An introduction using R, Second Edition, Chapman and Hall/CRC, 2016.