THESE DE DOCTORAT DE L'ETABLISSEMENT UNIVERSITÉ BOURGOGNE FRANCHE-COMTÉ

PREPAREE A UNIVERSITÉ DE BOURGOGNE

Ecole doctorale n°37

Sciences Pour l'Ingénieur et Microtechniques (SPIM)

Doctorat de Informatique

Par

Muhammad Arslan

Understanding Human Behaviors in Dynamic Building Environments

(Comprendre les comportements humains dans les environnements dynamiques de construction)

Thèse présentée et soutenue à Dijon, le 12/02/2020

Composition du Jury:

M. Christophe CRUZ	Maître de Conférences LIB - EA 7534, UBFC, Dijon, France	Directeur de thèse
M. Dominique GINHAC	Professeur ImViA - EA 7535, UBFC, Dijon, France	Codirecteur de thèse
M. Christophe CLARAMUNT	Professeur Naval Academy Research Institute Brest, France	Rapporteur
M. Jérôme GENSEL	Professeur Laboratoire d'Informatique de Grenoble, France	Rapporteur
M. Frank BOOCHS	Professeur Institute i3mainz, University of Applied Sciences, Germany	Examinateur
M. Eric LECLERCQ	Maître de Conférences LIB - EA 7534, Dijon, France	Examinateur

(Sciences Pour l'Ingénieur et Microtechniques)
(Engineering Sciences and Microtechnologies)
ED 37

Titre: Comprendre les comportements humains dans les environnements dynamiques de construction

Mots clés: Bâtiments, occupants, gestion de la sécurité, trajectoires spatio-temporelles, information contextuelle

Résumé: Dans un bâtiment, la présence ainsi que les mouvements des occupants sont fondamentaux et indispensables à la compréhension de tout type de comportement. Malgré de nombreuses recherches visant à modéliser les comportements dynamiques des occupants dans des bâtiments, la compréhension de ces comportements en intégrant les informations contextuelles liées à l'évolution de l'environnement bâti n'a toujours pas été suffisamment explorée. L'évolution de l'environnement du bâtiment affecte les mouvements et la présence des occupants à l'intérieur de l'installation, ce qui dégrade le processus de déduction de la précision de leurs activités en fonction du contexte. Désormais, les informations sur les bâtiments en évolution nécessitent d'être cartographiées avec les mouvements des occupants pour une meilleure compréhension de leurs comportements en évolution. Pour combler cette lacune en matière de recherche, une méthode appelée 'Occupant Behaviors in Dynamic Environments' (OBiDE) a été conçue pour fournir un modèle afin de mieux comprendre les comportements des occupants en traçant la dynamique des positions dans les bâtiments. Le modèle proposé fournit une base de connaissances centralisée contenant les mouvements des occupants avec des informations contextuelles historisées pertinentes de l'environnement du bâtiment afin d'étudier les comportements des occupants pour différentes applications de gestion de la sécurité.

Title: Understanding human behaviors in dynamic building environments

Keywords: Buildings, occupants, safety management, spatio-temporal trajectories, contextual information

Abstract: Occupants' movements and presence are fundamental and the pre-requisites for any type of occupant behaviors' understanding. Despite many research efforts to model dynamic behaviors of building occupants, the understanding of their behaviors by incorporating the contextual information linked to the evolution of the building environment is still not adequately explored. The evolution of the building environment affects the occupants' movements and presence inside the facility which ultimately degrades the process of inferring their accurate activities based on the location context. Henceforth, the evolving building information is required to be mapped with occupant movements for an improved understanding of their changing behaviors. To fill this research gap, a framework named 'Occupant Behaviors in Dynamic Environments' (OBiDE) is designed for providing a 'blueprint map' to better understand the occupant behaviors by tracking the dynamicity of building locations. The proposed framework provides a centralized knowledge base that holds the movements of occupants with relevant historicized contextual information of the building environment to study occupant behaviors for different safety management applications.

UNIVERSITE BOURGOGNE FRANCHE-COMTE

Laboratoire d'Informatique de Bourgogne (LIB) EA 7534, Univ. Bourgogne Franche-Comte, Batiment I3M, 64 rue Sully, Dijon 21000, France

Acknowledgements

First, I would like to thank God for giving me many blessings, energy, and patience to do this study.

I wish to convey my profound and sincere gratitude to my research director Dr. Christophe Cruz for his valuable and insightful guidance, and continuous support throughout my research. During my stay, he ensured that my work environment remains the best that one can wish for. Whether it’s the weekend, the holiday season, or the busy week, he was always there to guide me to fulfill the requirements for conference and journal publications. I accomplished a lot from his vast professional vision and scientific wisdom, and I couldn’t ask more in my director.

I am very thankful and indebted to my co-director Prof. Dominique Ginhac for trusting me, my work and always motivating me with his positive feedback. He was always there to support my trips to attend conferences and a summer school. His ability to sense the critical situations occurring during my stay, understanding the unsaid things and proactively favoring me, makes him an amazing person.

I am grateful to the French government for the research scholarship which helped me to complete my doctoral study uninterruptedly. Also, I am thankful to Université de Bourgogne (uB), especially Prof. Olivier Togni, Mme. Christiane Belin, Mme. Fabienne Blaise, Mme. Gaelle Dordain, Mme. Jessy Delpierre and M. Fabien George for taking care of all the administrative matters.

I am also indebted to Engr. Orval Touitou who helped me a lot in prototyping for this research. His positive attitude, and willingness to support the implementation requirements were tremendous. Lastly, but not the least I want to convey my special thanks to my family and friends. Meriem, Thomas and Roberto for helping me from the very 1st day when I arrived in Dijon. Yousaf, Raabid, Kibrom and Adrian for being the best group which is always fun and relaxing for me to be a part of it. Nérée, Cyril and Florian for helping me to understand the French environment, and for being good friends.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ACF	Autocorrelation function
AEC	Architecture, Engineering, and Construction
AIC	Akaike Information Criterion
API	Application Programming Interface
APs	Access Points
BBS	Behavior-Based Safety
BIM	Building Information Modeling
BLE	Bluetooth Low Energy
BLS	Bureau of Labor Statistics
CAD	Computer-Aided Design
DCT	Dublin Core Terms
DNAS	Drivers, Needs, Actions and Systems
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
GUI	Graphical User Interface
H&S	Health and Safety
HMMs	Hidden Markov Models
HSE	Health and Safety Executive
HVAC	Heating, Ventilation and Cooling
IFC	Industry Foundation Classes
IoT	Internet of Things
IPS	Indoor Positioning System
JHA	Job Hazard Analysis
JSON	JavaScript Object Notation
LIDAR	Laser Imaging Detection and Ranging
LoD	Level of Detail
LOD	Linked Open Data
LOS	Line-of-Sight
MCs	Markov Chains
NBIMS	National Building Information Modeling Standard
OBiDE	Occupant Behaviors in Dynamic Environments
OSM	OpenStreetMap
POI	Point of Interest
RFID	Radio-Frequency Identification
ROI	Region of Interest
RSSI	Received Signal Strength Indicator
STriDE	Semantic Trajectories for Dynamic Environments
VIDEWS	Visualizing Intrusions in Dynamic Environments

WoTAS	Worker Trajectory Analysis System
XML	Extensible Markup Language

Extended Summary (in French)

De nombreuses études ont été menées au cours des dernières décennies pour modéliser les comportements humains de manière stochastique afin de mieux comprendre leurs activités pour différentes applications de gestion technique des bâtiments. Malgré de nombreux efforts de recherche pour modéliser les comportements dynamiques des humains, il existe encore un fossé dans la compréhension de leurs comportements dans le contexte d'environnements dynamiques de construction. Les données contextuelles liées aux changements dans des environnements dynamiques évoluent souvent en termes de position, de taille, de propriétés et de relations avec l'environnement. Ces changements affectent les mouvements et le comportement des occupants à l'intérieur de l'installation, ce qui dégrade le processus de déduction de leurs activités précises en fonction de la localisation dans le bâti. Désormais, les informations sur les bâtiments en évolution doivent être cartographiées ainsi que les mouvements des occupants pour une meilleure compréhension de leurs comportements en évolution. Pour combler cette lacune en matière de recherche, une plateforme appelée OBiDE (Occupant Behaviors in Dynamic Environments) a été conçue. Elle fournit un plan détaillé permettant d'intégrer le modèle DNAS (Drivers, Needs, Actions et Systems) capable de modéliser les comportements des occupants, au modèle STriDE (Semantic Trajectories in Dynamic Environments), capable de capturer la dynamique des bâtiments pour une meilleure compréhension des comportements. La plateforme proposée étend le modèle DNAS en produisant une base de connaissances centralisée contenant les mouvements avec des informations contextuelles historisées pertinentes de l'environnement du bâtiment afin d'étudier les comportements des occupants pour différentes applications de gestion d'installations. La plateforme intégrée OBiDE permet d'extraire les états comportementaux des occupants et a été utilisée pour améliorer la sécurité des travailleurs sur les chantiers de construction. Le choix des environnements de chantier comme cas d'utilisation a pour objectif de montrer le fonctionnement et la faisabilité de la preuve de concept de la plateforme proposée. Ce choix permet de montrer la prise en compte de l'évolution de l'information contextuelle dans les bâtiments et les sites de construction connus pour être parmi les exemples les plus représentatifs des environnements dynamiques.

Les travailleurs du bâtiment sont plus souvent exposés aux environnements dynamiques rudes comportant des risques élevés pour la sécurité par rapport aux autres secteurs d`activité. Les risques pour la sécurité découlent de l'incertitude des comportements des travailleurs sur les sites, car leurs actions risquent de dévier d'un plan prédéfini en raison des environnements dynamiques. Les comportements incertains des travailleurs peuvent potentiellement créer des risques pour la sécurité, entraînant des situations dangereuses. Indépendamment de la disponibilité de nombreux systèmes de surveillance pour la sécurité dans le secteur de la construction, le taux de décès reste élevé. Un examen plus approfondi des études récentes révèle que la plupart des accidents dans le secteur de la construction sont imputables aux comportements dangereux liés aux mouvements des travailleurs qui entraînent souvent des collisions graves avec des objets et des machines du site. Les accidents mineurs (mouvements brusques et rotations ou intrusions) doivent être analysés pour éviter les blessures non létales qui pourraient éventuellement réduire le taux d'accidents. Les mouvements dangereux (mouvements brusques et rotations ou intrusions) non seulement affectent les travailleurs non autorisés et victimes d'accidents, mais peuvent également interrompre ou même nuire à d'autres travailleurs dans les zones à risque. Les accidents évités de justesse sur les sites sont souvent négligés, car les méthodes actuelles d'évaluation de la sécurité dans le domaine de la construction sont principalement axées sur les conséquences visibles telles que les blessures et les décès. La principale raison de cette négligence est la complexité de la reconnaissance des accidents mineurs et le manque de production de rapports pendant les travaux de construction. Pour résoudre ce problème, une formation sur la sécurité basée sur le comportement (BBS) est généralement dispensée pour améliorer l'attitude des travailleurs vis-à-vis de l'organisation pour une meilleure culture de la sécurité sur les sites. Indépendamment des nombreuses implémentations réussies de la formation BBS dans de nombreux projets de construction, ces formations ont des limitations telles que: 1) incapable de dépasser les limites imposées aux travailleurs pour identifier toutes les zones dangereuses environnantes; et 2) principalement de se fier aux inspections effectuées par des superviseurs de la sécurité bien formés et de ne pas fournir de rétroaction rapide pour changer les comportements dangereux des travailleurs. Pour surmonter les limites de la formation BBS, l’un des moyens les plus efficaces pour réduire le nombre de comportements à risque et réduire les risques d’accidents mineurs est de surveiller les comportements des travailleurs grâce à l’étude de leurs mouvements spatio-temporels capturés par un système de géolocalisation.

Ainsi, pour comprendre les mouvements spatio-temporels des travailleurs dans leur environnement de construction impliquant le déplacement et la modification d’objets, un système appelé ‘WoTAS’ (Worker Trajectory Analysis System) est proposé. Premièrement, un sous-système de collecte de données et de prétraitement de trajectoire basé sur des balises Bluetooth Low Energy (BLE) est conçu pour extraire les caractéristiques multifacettes des trajectoires. Deuxièmement, après avoir collecté le jeu de données de trajectoires, les zones de travail sont identifiées ainsi que la durée. Ces zones de travail aideront à reconnaître les régions importantes dans le bâtiment pour la catégorisation des mouvements de travailleurs. L’objectif de l’extraction des emplacements dans les trajectoires est d’identifier les régions d’intérêt (ROI) et les positions d’intérêt (POI) associées. Les ROI sont les zones les plus larges du bâtiment, comprenant plusieurs points d'intérêt géographiques (par exemple une voie de circulation extérieure, une salle de stockage, un bureau, etc.) étiquetés comme des salles de notre modèle. Étant donné que les chantiers de construction sont divisés en différentes zones, telles que la zone de matériaux, la zone de dépôt, la zone de chargement, la zone de travail, etc., les superviseurs des bâtiments facilitent le suivi et la gestion du site. Le système WoTAS développé a identifié plusieurs régions de séjour dans les trajectoires des utilisateurs qui correspondent à différentes ROI. Pour faire correspondre les trajectoires traitées avec les ROI, des jointures spatiales sont effectuées. Cependant, avant d’acquérir les données de trajectoire des occupants, une analyse préliminaire des données de mesure des balises (BLE) a été réalisée afin d’obtenir une meilleure précision dans l’identification des emplacements des bâtiments. Troisièmement, pour que les informations sémantiques aident à mieux comprendre les mouvements de travailleurs en exploitant des données contextuelles liés à l’environnement de la construction, le modèle STRIDE basé sur une ontologie est proposé permettant de suivre l’évolution des objets en capturant les trajectoires sémantiques. Pour extraire des informations des trajectoires sémantiques, le modèle de Markov à états cachés (HMM) est l’une des approches probabilistes présentes dans la littérature pour décrire le comportement. À l'aide du modèle HMM, un ensemble de trajectoires appartenant à une région de séjour est analysé en catégorisant les mouvements des travailleurs dans trois états différents. En fin de compte, le résultat de l'algorithme de Viterbi est affiché à l'aide d'un outil BIM afin d'identifier les emplacements à haut risque les plus probables impliquant des mouvements et des rotations de travailleurs dangereux. L'approche BIM est choisie car la littérature présente cette approche comme très pertinente. Elle est préférée aux approches CAD 3D traditionnelles, car elle constitue un moyen efficace de gestion des informations pendant le cycle de vie du bâtiment et de l'analyse de la sécurité. Surtout, l'approche BIM devient la norme de référence dans l'industrie de la construction dans de nombreux pays.

Le deuxième scénario d'application de la plateforme OBiDE est appelé VIDEWS (Visualizing Intrusions in Dynamic Environments). Il propose d’effectuer la détection des intrusions dans les bâtiments en 3 étapes. La première étape permet d’acquérir les mouvements de travailleurs à l’aide de balises BLE (Bluetooth Low Energy). Ensuite, il s’agit d’effectuer le prétraitement des données acquises et enrichir sémantiquement les trajectoires spatio-temporelles des travailleurs en utilisant les informations contextuelles pertinentes mises à jour du bâtiment. Enfin, la troisième étape utilise l’approche BIM (Building Information Modeling) pour visualiser les emplacements des bâtiments où des intrusions ont eu lieu.

Les systèmes développés (WoTAS et VIDEWS) réalisent l'intégration de plusieurs sous-systèmes allant de l'étape d'acquisition des données jusqu’à la visualisation des comportements des travailleurs dangereux. Les résultats obtenus peuvent servir de base à des recherches futures sur les comportements avancés des travailleurs liés aux mouvements dans des environnements dynamiques en surmontant les contraintes spatio-temporelles de la gestion des données. Les responsables de la santé et de la sécurité peuvent utiliser une visualisation générée pour les outils BIM décrivant les comportements des utilisateurs, afin d'améliorer les stratégies d'intervention en matière de gestion de la sécurité en analysant les trajectoires sémantiques et en identifiant les emplacements des bâtiments. Les systèmes ainsi développés aideront les responsables de la sécurité à surveiller et à contrôler les activités des bâtiments à distance dans des environnements dynamiques. Comprendre les mouvements de travailleurs améliorera la gestion de la sécurité dans les opérations de construction quotidiennes. À terme, ces systèmes contribueront à réduire les risques d'incidents évités de justesse sur des sites susceptibles de provoquer des accidents graves.

Chapter 1 – Introduction

Outline

The work presented in this thesis focuses on understanding spatio-temporal occupant behaviors in dynamic building environments. Occupants' movements and presence are fundamental and the prerequisites for any type of occupant behaviors' understanding. The contextual information linked to locations in dynamic building environments changes often over time in terms of position, size, properties, and relationships with the environment. This changing building environment affects the occupants' movements and presence inside the facility which ultimately degrades the process of inferring their accurate activities based on the location context. Henceforth, the evolving building information is required to be mapped with occupant movements for an improved understanding of their changing behaviors. To work on this requirement, initially, the relevant literature on occupant behaviors is reviewed. First, occupant behaviors and their different types are defined. Then, it describes the methods to collect sensory data related to occupants as well as building environment data to study occupant behaviors using a building context. A brief introduction of the Building Information Modeling (BIM) approach is presented to describe its application for extracting building data to study occupant behaviors along with real-time sensory data. Later, a review of the existing literature has identified that BIM and sensor-based integrated solutions have been used extensively for a broad range of applications for studying occupant-building behavioral interactions. These integrated systems capture the dynamicity of occupants using real-time sensory data for their behavioral understanding. However, the scope of this research is centered on dynamic building environments. For this, the concept of a dynamic environment is explained with the help of a construction site scenario. Later, a brief introduction to the occurrences of fatal accidents on construction sites is presented with the help of OSHA statistics. The H&S statistics highlighted that most of the accidents which have occurred in the construction sites are because of the unsafe movement behaviors of the construction workforce (workers and machinery). Further review of the related literature has identified that the dynamicity of the site environment is not adequately incorporated during the modeling of the worker behaviors for safety management applications. Based on this research gap, research questions were formulated, and a solution is presented using an 'Occupant Behaviors in Dynamic Environments' (OBiDE) framework to study occupant movements in dynamic building environments with the help of two different construction applications. In the end, an organization of this work is mentioned.

1.1 Background and Motivation

Humans perceive, move and perform actions in the physical world, and make interactions with the objects, the other humans, and the environment to achieve their goals for attaining the desired level of satisfaction (Wagner et al., 2018; Chen et al., 2015). A series of entangled mental activities from raw instincts to high-level reasoning initiates human choices, actions, and interactions with the physical environment (Chen et al., 2015). Responding to human actions when they have been executed allows analysts to examine their complete effects and have a comprehensive picture of the environment. But the outcomes of their performed actions may create serious problems for other humans and the environment. For preventing such undesirable situations, a more dynamic and real-time understanding is required while their actions are still evolving in the physical environment (Chen et al., 2015). Understanding the humans implies inferring the goals and the motivations using their actions for recognizing their impact on the environment and thus to agree which reaction is more suitable for effective building management (Hong et al., 2015).

More formally, human behaviors can be defined using a building context as; observable actions or reactions of humans in response to external or internal stimuli (Wagner et al., 2018; Hong et al., 2015; Yan et al., 2015). These actions or reactions can be categorized into four main types which are; physiological adjustments (e.g. sweating, shivering, etc.), individual adjustments (e.g. selection of clothes, using earplugs, etc.), environmental adjustments (HVAC adjustment, window opening or closing, etc.) and spatial adjustments (moving from one building facility to another, etc.) (Yan et al., 2015). In other words, behaviors are the interactions (leaving or entering a room, visual and thermal indoor conditions adjusting using windows or blinds, doors, etc.) of building users which can be categorized into different movements by modeling primarily their presence or actions` data with the environment (building, its systems and appliances) that impact on the building performance (heating or cooling, indoor air quality, energy, comfort, etc.) during a building lifecycle (Wagner et al., 2018; Yan et al., 2015). Thereby, a human interaction that results in changing a building state (presence or absence in case of occupancy monitoring) or no interaction leaving the present state of a building unchanged are both facets of human behaviors (Wagner et al., 2018).

Over the last decade, human behavior understanding has attracted the strong interest of the researchers (Wagner et al., 2018) for better monitoring and controlling the facilities. This massive interest in behavior understanding is because of its numerous potential applications which can be divided into three major domains as shown in Fig. 1.1. To understand the human behaviors for different applications, in the existing literature (Wagner et al., 2018; Chen et al., 2015; Hong et al., 2015; Yan et al., 2015), an extensive range of different types of sensors (wired and wireless) for monitoring humans and the environment are present to acquire rich information for modeling human behaviors and their interactions (energy consumption, etc.) with the buildings. The sensors are often low-cost nodes and typically connected to the internet for data transmission. After collecting the required sensory data, this data needs to be mapped with the corresponding building information for analyzing the human behaviors using a building context (Z. Yan, 2011;

Arslan et al., 2018; 2019a). To employ the building information, building information model (BIM)-based platforms are utilized and preferred over traditional 3D CAD-based systems in the Architecture, Engineering, and Construction (AEC) industry.

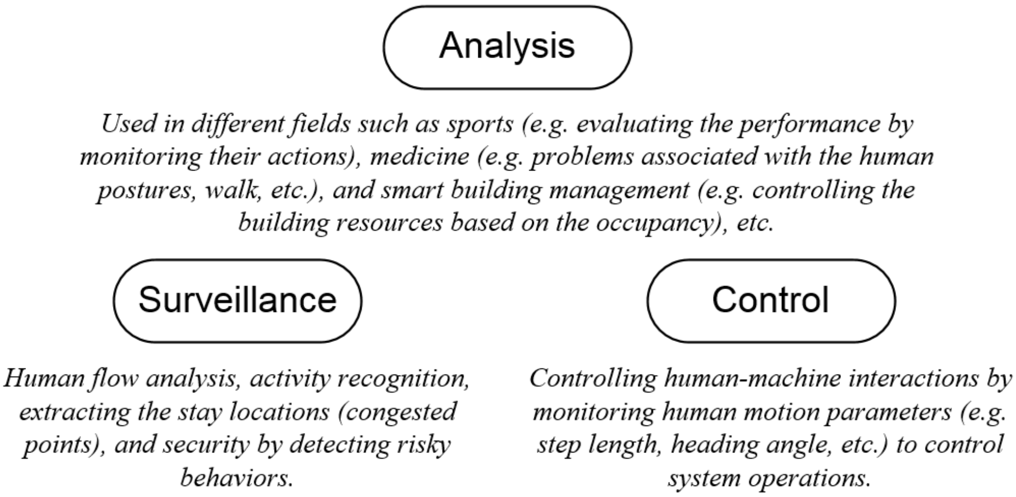

Fig. 1-1 Applications of understanding human behaviors (Wagner et al., 2018; Chen et al., 2015; Hong et al., 2015; Yan et al., 2015)

(Riaz et al., 2014). A BIM model is a digital representation of the physical and functional characteristics of a facility providing a source of shared knowledge of facility information for the building managers to use and maintain information throughout its lifecycle (NBS-US, 2019). The adoption rate of BIM in the US AEC industry has risen from 17% in 2007 to 71% in 2012. Whereas, in the UK AEC industry, it has improved to 74% in 2018 (Maltese et al., 2017). The application of BIM is presently going through tremendous growth in enhancing building operations for better planning and management by enabling people to work collaboratively and cost-effectively (Maltese et al., 2017). Every building component defined in a BIM model is a solid object containing the geographic information as well as the alphanumeric properties of an object (NBS-US, 2019; Liu et al., 2019). However, BIM lacks the functionality of assessing the current state of the building environment in real-time for building management (Stojanovic et al., 2018). Here, a state of a building object (e.g. a location) refers to its environmental data (e.g. temperature, humidity, etc.), the status of people inside the location (e.g. tracking, occupancy monitoring and identifying their actions), analyzing the utilization rate of building equipment or location, and performing localization and preventive maintenance of building objects for different BIM-based application scenarios (Tang et al., 2019). BIM data needs to be constantly updated with the current state of the building environment to accommodate the real-time monitoring of building locations for building management (Stojanovic et al., 2018; Tang et al., 2019).

To incorporate the sensory data related to a building environment into a BIM model for studying human behaviors in the context of a building, wireless sensor technology has gained great importance in real-time monitoring of the buildings (Yan et al., 2015). Integration of BIM data with real-time sensory data acquired from the Internet of things (IoT), Radio Frequency

Identification (RFID), 3D scanning, GPS, etc. has provided us a centralized digital building model to study different types of human behaviors for building management (Tang et al., 2019). However, this process of understanding human behaviors gets more complicated when the dynamicity of the environments gets incorporated into the behavior extraction process (Cruz et al., 2017). Our physical environments are dynamic environments that evolve in terms of geometry and contextual information. Here, a context refers to any information about physical space, time and environment utilized for categorizing the situation of humans in buildings (Yan et al., 2015). The best example of a dynamic environment to study different worker behaviors in the context of an evolving building environment is the construction site. New walls and infrastructure supports are often added on construction sites, while others are detached. This opens more challenges to keep track of the changes in the attributes of the locations which evolve for understanding behaviors of the construction workforce (workers and machinery) (Pradhananga and Teizer, 2013). For example, a storage zone in the building is now a work zone having a different functionality (i.e. a semantic change). Moreover, the dimensions of the dumping zone are changed because of the construction of a new wall on a site (i.e. a spatial change). Alike situations occur very often on the construction sites. Such changes in the purpose or the position of the locations in the buildings or on construction sites will result in different behaviors of workers (Yan et al., 2015). The updated context about the locations along with the previous contextual information is required to be captured for studying the behaviors of the building objects including the workers in detail concerning the changes occurred in the building environment. The historicization of building information along with worker location data will help building managers for conducting 'cause and effect' analysis (Arslan et al., 2019b). The resulted analysis can be helpful for a broad range of applications in the area of site monitoring (Volk et al., 2014; Costin et al., 2015) relating to the management of a changing environment which involves moving objects as well as evolving building geometries such as construction planning, risk and hazard assessment, clash detection, construction logistics, building automation, spatial analysis, etc. (Arslan et al., 2019b; 2019c). However, the scope of this research is kept limited to studying worker behaviors on construction sites for safety management. The reason for selecting this area is because the higher level of uncertainty exists in the worker behaviors as their activities are likely to deviate from a predefined plan because of the evolving construction environment which raises serious safety concerns. The uncertain worker behaviors if not analyzed with the evolving site environment information can potentially create safety risks resulting in hazardous situations (Costin et al., 2015).

1.2 Research Gap

According to Heinrich`s triangle (see Fig. 1.2) that fatalities on construction sites typically occur after several non-fatal injuries (Bellamy et al., 2015). Similarly, non-fatal injuries are the successor of multiple first-aid injuries. Also, first-aid injuries are the result of several near-misses. A near-miss (Zhou et al., 2019) consists of two parts; 1) 'near' which means 'close' (e.g. being in proximity to success or failure), and 2) 'miss' which means 'lose' (e.g. losing a chance to reach a goal or to avoid failure). In literature, there exist many factors of having near-miss incidents which can potentially lead to accidents on construction sites.

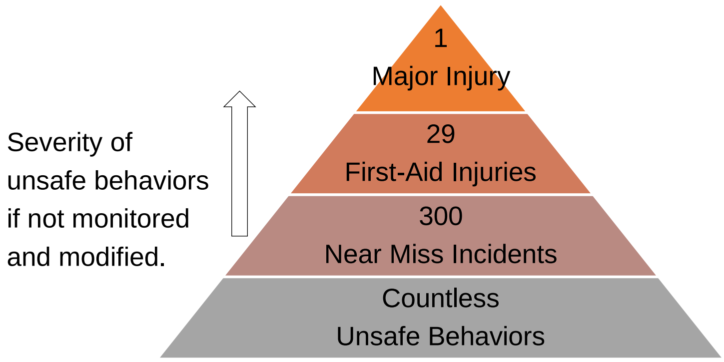

Fig. 1-2 Modified Heinrich's triangle (Bellamy et al., 2015)

A few of them are; 1) unsafe worker behavior, 2) unsafe construction site environment, 3) incidents with property loss, and 4) incidents with possible damage to the site environment (Zhou et al., 2019; Choudhry et al., 2008). The near-misses (see Fig. 1.3) develop into accidents due to the presence of opportunity factors. The opportunity factors are beyond normal control and if opportunity factors exist, an accident will happen with serious consequences (Zhou et al., 2019). Ultimately, the reasons behind near-misses (unsafe acts which potentially cause a fatality) are risk-taking (unsafe) behaviors. One of the most effective ways to reduce the occurrences of risky behaviors for reducing the chances of near-misses which eventually lead to accidents is by monitoring the worker behaviors in real-time (Bellamy et al., 2015).

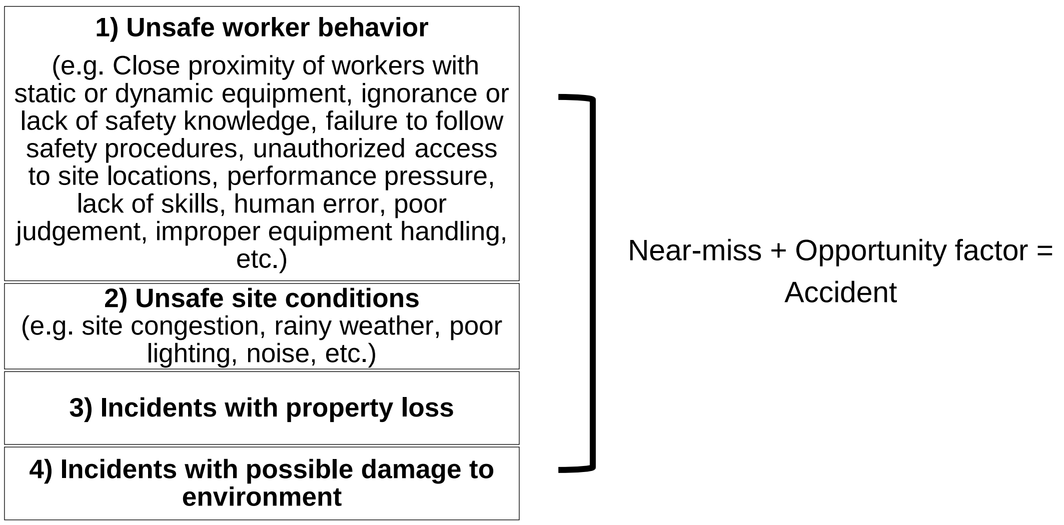

Fig. 1-3 Near-misses leading to accidents (Zhou et al., 2019)

To choose different types of worker behaviors that need to be studied in the dynamic environment context that causes severe safety risks on sites, the most updated available statistics of the fatal occupational injuries on sites were reviewed. According to the Bureau of Labor Statistics (BLS), in 2017 out of 5,147 fatal occupational injuries, 971 were recorded from the U.S. construction industry (BLS, 2017) (see Fig. 1.4). Out of 971, a major number of fatalities were resulted from falling from heights (i.e. 39.2%) and struck-by-object incidents (i.e. 8.2%). Further analysis of the existing studies (Zhou et al., 2019; Wu et al., 2010; Yang et al., 2019) to explore different unsafe behaviors related to falling from heights and struck-by-objects revealed that majority of the fatal injuries were caused by the "unsafe movement-related" worker behaviors which resulted in serious accidents on sites. For example, a) limited spatial awareness of the operating construction machinery involving sharp movements and rotations within the workers` proximity due to blind spots and surrounding noise, and b) an unauthorized staying in or stepping into the hazardous areas without realizing the potential dangers associated with the locations and are very common type of near-miss incidents (see Fig. 1.5) have led to hazardous situations on sites in previous years.

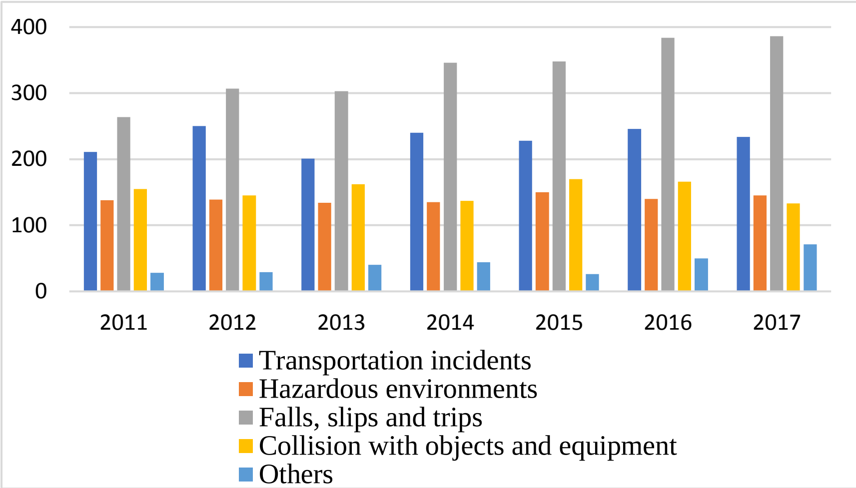

Fig. 1-4 Fatal injuries in the U.S. construction industry (BLS, 2017)

Existing literature (Teizer et al., 2015; Cheng et al., 2011; Costin et al., 2015; Zhang et al., 2015) consists of numerous solutions based on location acquisition sensing technologies for tracking the movements (risky movements and unauthorized access) of construction workforce (workers and machinery) in real-time to generate timely alerts for building and safety managers to take necessary actions in case of unsafe behavior detection.

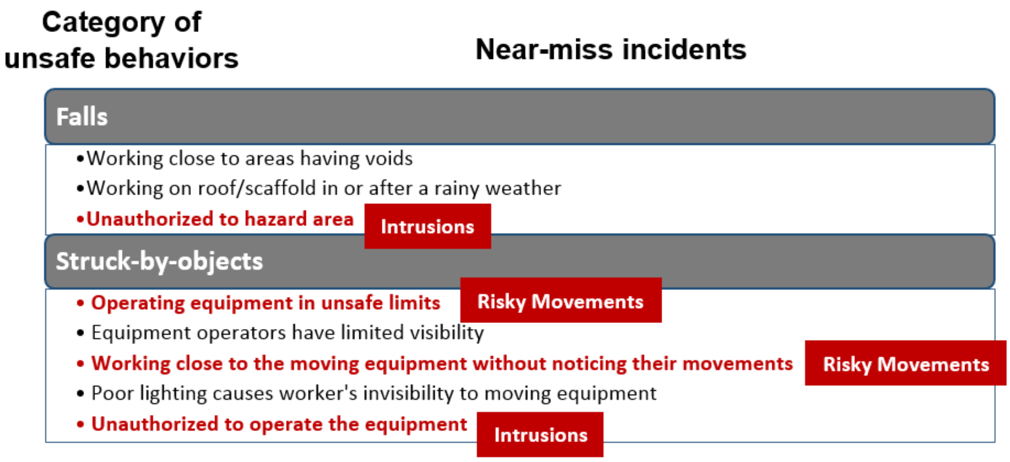

Fig. 1-5 Risky movements and intrusions, the leading causes of near-misses (Wu et al., 2010)

1.3 Research Objectives

To enrich the location trajectories of workers with the contextual information of the building, building data is typically extracted using an OSM or a BIM-based data file (Arslan et al., 2018). However, a BIM-based method is preferred and widely adopted for sensor data integration (Tang et al., 2019). BIM and sensor-based approaches (Tang et al., 2019) have been used extensively for designing different applications of construction site monitoring and management for studying dynamic worker behaviors (e.g. occupancy, movements, activities, etc.) using the up-to-date site information. However, the studies of dynamic worker behaviors in the context of an evolving construction site environment by keeping its historicization of the changes occurred during its lifecycle are not yet adequately explored. In the existing BIM and sensor-based solutions, once the building model is modified by incorporating the spatial and contextual (i.e. alphanumeric properties of locations) changes by the AEC managers during the construction stage, the previous information of building is lost (Stojanovic et al., 2018). This case is not only limited to the construction phase of buildings but also applies to the rest of the building phases where the building undergoes spatial evolutions less often but more frequent changes in the contextual information linked to building locations. It is also not feasible to manage different versions of BIM models after every modification in buildings as modifications in the building infrastructure and associated contextual information are countless during a building lifecycle (Jaly-Zada et al., 2015). The historicized information of building locations involving the spatial and contextual evolutions is important to keep so that different behaviors of building entities (workers and building equipment) can be studied in detail with respect to the type of changes that occurred in the building environment.

The historicization of building information will help AEC managers for conducting ‘cause and effect’ analysis (Arslan et al., 2019a; 2019b). For example, extracting the type of change occurred in the building infrastructure which resulted in an accident. Moreover, resulted worker behaviors from the analysis can be used for different construction and built environment applications such as; construction resource monitoring for improved safety, managing building spaces based on their utilization, implementing worker access control systems, etc. Failure in understanding worker behaviors because of inadequate integration of all relevant contextual factors associated with the workers and the evolving facilities can result into serious financial and management crises such as under-utilization of the site facilities, decreased productivity due to poor environmental conditions in buildings, increased energy usage, and safety hazards (Wagner et al., 2018; Chen et al., 2015; Yan et al., 2015). On the contrary, if the worker behaviors are modeled and predicted effectively by incorporating all the possible contextual factors (social-personal, economic, etc.) and the dynamicity of the site environments which may influence worker behaviors will lead to an increased physical comfort, enhanced safety at work and improved work performance of the workers while keeping the level of site facilities` resources to the optimum (Hong et al., 2015).

1.3 Research Objectives

Existing literature encompasses many studies based on predominately BIM and sensor integrated solutions for constructing worker behavior extraction systems which can help construction and facility managers in decision-making for building and site safety management by stochastically modeling the dynamic behaviors of occupants by incorporating the random variations in their behaviors over time (Tang et al., 2019). However, the dynamicity of the building environments (i.e. evolving contextual information of a building) is still not yet adequately incorporated into BIM and sensor-based behavior extraction systems to provide the historicization of information concerning worker movements mapped with the up-to-date building information. To address this research gap, two main research questions (Q1 and Q2) were identified to be discussed in this dissertation;

(Q1) How to establish a high-level understanding of the human behaviors in a dynamic environment using a generalized framework which should be constructed using the fundamental components of the state-of-the-art systems for human behavioral understanding?

(Q2) How to analyze the human behaviors in dynamic environments such as construction sites using the historicized information of human movements mapped with the related building contextual information for safety management applications?

1.4 Contributions

To address the abovementioned research questions, the contributions of this thesis were made (see Fig. 1.6) based on the research method (see Fig. 1.7) as described below.

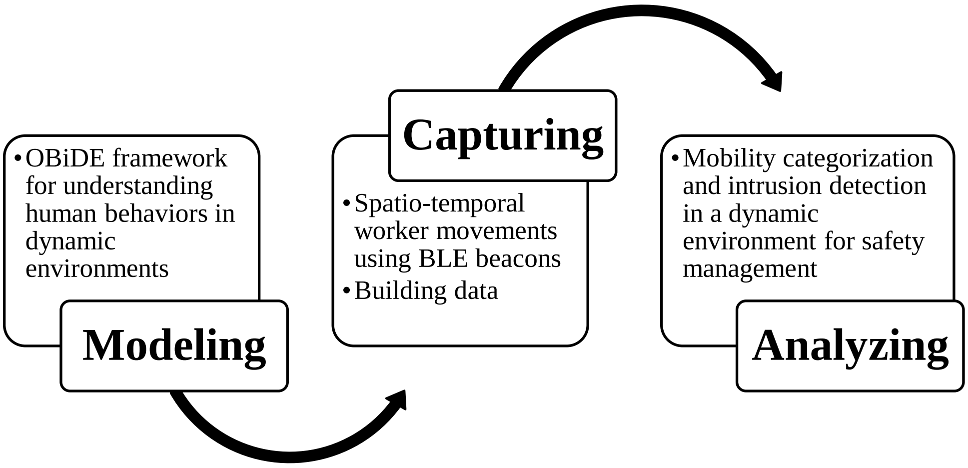

Fig. 1-6 A generic view of thesis contributions

For human behavioral understanding, initially existing human behavior extraction systems were reviewed from the literature (Wagner et al., 2018; Chen et al., 2015; Hong et al., 2015; Yan et al., 2015). Most of the systems were exclusively designed and implemented for modeling specific human behaviors such as opening or closing of the windows by humans in buildings or studying the mobility trends within the facilities. However, a DNAS (Drivers, Needs, Actions, and Systems) framework was chosen from the literature (Hong et al., 2015) as it simply describes human behaviors using four parameters which are; 1. the ‘drivers’, 2. the ‘needs’, 3. the ‘actions’ which building occupants perform to fulfill their needs and 4. the ‘systems’ with which the occupants interact to perform actions. Although, the application of the DNAS framework (Hong et al., 2015) is primarily explored for energy-related occupant behavior modeling in the literature. However, this framework is treated as a foundation in this dissertation to represent other types of human behaviors as it provides a basic ontology to represent human behaviors with the help of four core components as discussed above. After choosing a basic DNAS ontology to describe human behaviors, it was found in the literature review (Wagner et al., 2018) that every human behavior modeling process initiates after the tracking of the human movements and presence inside the facilities. In fact, human movements and presence are fundamental and the prerequisites for any type of human behaviors` understanding which tells whether a facility is occupied and the number of humans with a specific profile in a certain location (Yan et al., 2015).

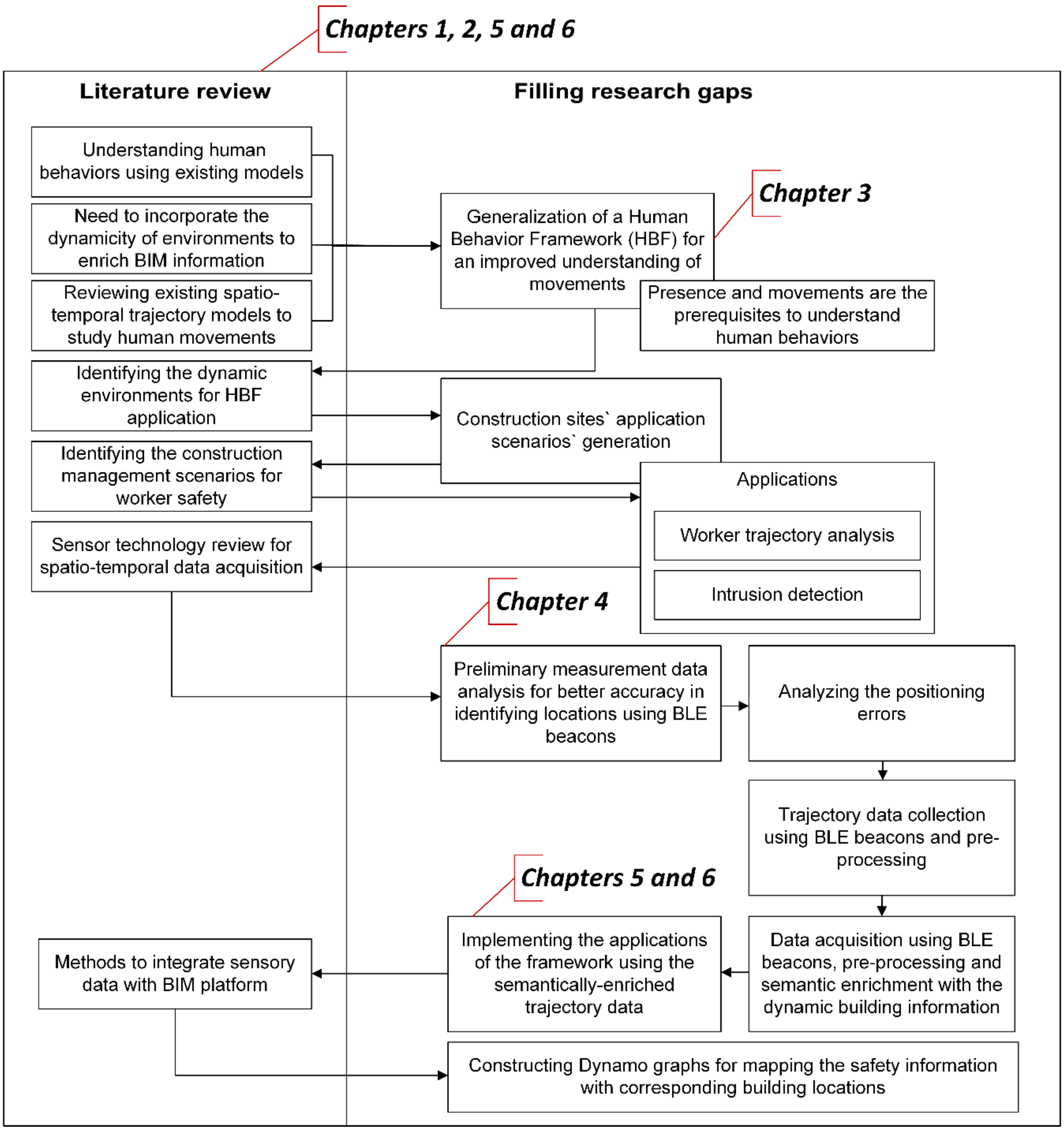

Fig. 1-7 Research method

Numerous studies (Wagner et al., 2018; Chen et al., 2015; Hong et al., 2015; Yan et al., 2015) have been conducted over the past few decades to model human behaviors stochastically for an improved understanding of their activities for different facility management applications. Despite many research efforts to model the dynamic behaviors of humans, there is still a gap exists in understanding their behaviors in the context of dynamic building environments. The contextual data linked to locations in dynamic environments often evolve in terms of position, size, properties, and relationships with the environment (Cruz et al., 2017). This changing building environment affects the occupants` movements and presence inside the facility which ultimately degrades the process of inferring their accurate activities based on the location context. Henceforth, the evolving building information is required to be mapped with occupant movements for an improved understanding of their changing behaviors. To fill this research gap, a framework named 'Occupant Behaviors in Dynamic Environments' (OBiDE) is designed (Arslan et al., 2019d; 2019e) for providing a 'blueprint map' to integrate existing DNAS (drivers, needs, actions and, systems) model (i.e. a scheme to model occupant behaviors) with the STriDE (Semantic Trajectories in Dynamic Environments) data model to include the dynamicity of building locations for an improved understanding of human behaviors. The proposed framework extends the usability of DNAS by providing a centralized knowledge base that holds the movements of humans with relevant historicized contextual information of the building environment to study human behaviors for different facility management applications. After constructing an integrated framework i.e. OBiDE for extracting the behavioral states of the human behaviors, the framework is utilized for enhancing worker safety on the construction sites. The reason for choosing the use-cases related to the construction site environment is to show the proof-of-concept working and feasibility of the proposed framework which is designed to hold evolving contextual information regarding buildings and the construction sites are the best example of such dynamic environments.

The construction workers are more often exposed to the harsh dynamic environments involving high safety risks as compared to the other work sectors (Li et al., 2018). The safety risks arise from the uncertainty in the worker behaviors on sites as their actions are likely to deviate from a predefined plan because of the dynamic environments (Teizer et al., 2015). The uncertain worker behaviors can potentially create safety risks resulting in hazardous situations (Bellamy et al., 2015). Regardless of the availability of numerous safety monitoring systems (Li et al., 2018) for the construction industry, a rate of fatalities continues to be high (BLS, 2017). A closer look at the recent studies (Zhou et al., 2019; Wu et al., 2010; Yang et al., 2019) reveals that most of the construction fatalities are caused by the unsafe worker movement-related behaviors (sharp movements and rotations, or intrusions) which often result in serious collisions with site objects and machinery. The near-misses (sharp movements and rotations or intrusions) are required to be analyzed for avoiding non-fatal injuries that will eventually decrease the rate of fatal and non-fatal accidents on sites (Bellamy, 2015). The unsafe movements (sharp movements and rotations or intrusions) not only cause unauthorized workers to get affected and suffer from an accident but can also interrupt or even harm other workers in the hazardous site zones (Huang and Hinze, 2006). Near-miss incidents on sites are often overlooked as current construction safety assessment methods are mainly focused on visible consequences such as injuries and deaths (Heng et al., 2016). The major reason for getting neglected is because of the complexity in recognizing near-misses and providing feedback in real-time during construction operations (O'Neill et al., 2013). To solve this problem, Behavior-Based Safety (BBS) training (Guo et al., 2018) is usually conducted to improve the worker attitude towards organizing a better safety culture on sites. Regardless of many successful implementations of BBS training in many construction projects, they have limitations as well such as; 1) unable to overcome the worker limits to identify all surrounding hazardous zones, and 2) mostly relying on the physical inspections by well-trained safety supervisors and failing to provide quick feedback to change unsafe worker behaviors in real-time (Fang et al., 2018; Heng et al., 2016). To overcome the limitations of BBS training, one of the most effective ways to reduce the occurrences of risky behaviors for reducing the chances of near-misses which eventually lead to accidents is to monitor worker behaviors using their spatio-temporal movements captured using geo-localization systems in real-time (Teizer et al., 2015).

For understanding the spatio-temporal worker movements in dynamic construction environments that involve moving and changing objects, a system named 'WoTAS' (Worker Trajectory Analysis System) is proposed (Arslan et al., 2019b). First, a real-time Bluetooth Low Energy (BLE) beacons-based data collection and trajectory pre-processing subsystem is built for extracting multifaceted trajectory characteristics (Arslan et al., 2018). Second, after collecting the required dataset of trajectories, workers` stay regions are identified with their stay duration. The stay regions of the workers will help in recognizing the important regions in the building for categorizing the worker movements. The reason for extracting such locations in trajectories is to identify the Region of Interests (ROIs) and their associated Position of Interests (POIs). The ROIs are the wider areas of the building (e.g. work-zone237, etc.) which include multiple geographical POIs (e.g. outdoor pathway, storage room, office1, etc.) labeled as 'rooms' in our model as shown in the Fig. 1.8. The idea of splitting a building floor into a set of ROIs is taken from the Pradhananga and Teizer, 2013 research. As construction sites are divided into different zones such as material zone, dumping zone, loading zone, work zone, etc. by the building supervisors for easy and well-organized site monitoring and management. The developed WoTAS system has identified several stay regions in users` trajectories that correspond to different ROIs. For mapping the processed trajectories with the ROIs, spatial joins are performed (Z. Yan, 2011). However, before acquiring the trajectory data of occupants, preliminary measurement data analysis of beacon data was conducted for achieving better accuracy in identifying the building locations. Third, to enable the desired semantic insights for better understanding the underlying meaningful worker movements using the contextual data repositories related to the building environment, an ontology-based 'STriDE' (Semantic Trajectories for Dynamic Environments) model (Cruz et al., 2017) is applied which tracks the evolution of moving and changing building objects and outputs semantic trajectories. For extracting insights from the semantic trajectories, the Hidden Markov Model (HMM) is used (Rabiner, 1989) which is one of the probabilistic approaches present in the literature for describing the object behavior in time. Using the HMMs, a set of trajectories belonging to a stay region is analyzed by categorizing the worker movements into three different states. In the end, the output of the Viterbi algorithm (Rabiner, 1989) is visualized using a BIM model for identifying the most probable high-risk locations involving sharp worker movements and rotations. The BIM approach is chosen for generating visualizations because literature identifies it as a 'future IT solution' and preferred over traditional 3D CAD approaches as it is an efficient way of information management during the building lifecycle for safety analysis (Riaz et al., 2014). Above all, the BIM approach is becoming a construction industry standard in many countries (Riaz et al., 2014).

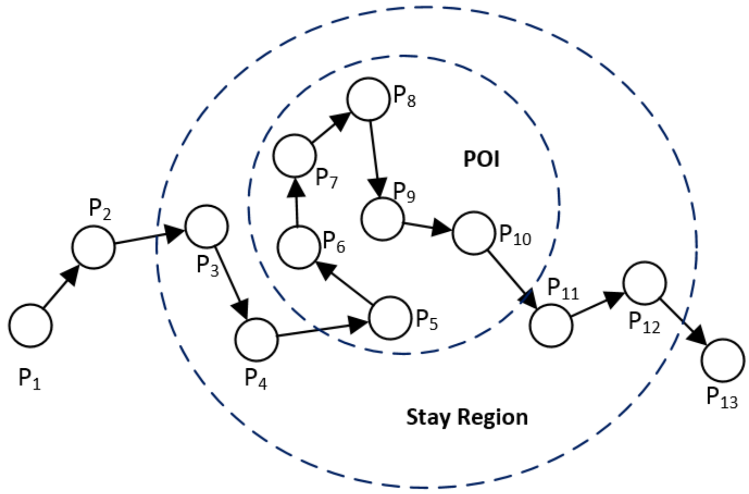

Fig. 1-8 A stay region (an ROI) and a POI

The second application scenario for an OBiDE framework named 'VIDEWS' (Visualizing Intrusions in Dynamic Environments) is proposed (Arslan et al., 2019c) for detecting intrusions in buildings. First, to acquire the worker movements using Bluetooth Low Energy (BLE) beacons. Second, to perform the pre-processing as well as semantically enriching the spatio-temporal worker trajectories using relevant updated contextual information of a building. Lastly, to visualize the building locations where intrusions have occurred using Building Information Modeling (BIM) approach.

The developed systems (WoTAS and VIDEWS) present the integration of systems from the data acquisition stage to visualizing the unsafe work behaviors that could serve as a foundation for future research in studying advanced movement-related worker behaviors in dynamic environments by overcoming the spatio-temporal data management challenges. Generated BIM-based visualization depicting user behaviors can be used by the H&S managers that can lead to improved safety management intervention strategies by analyzing semantic trajectories and identifying building locations involving high-risk worker movements to prevent accidents. The developed systems will help safety managers in monitoring and controlling building activities remotely in dynamic environments. Understanding worker movements will improve safety management in day-to-day building operations. Eventually, will also contribute to reducing the chances of near-miss incidents on sites that have the potential to cause serious accidents.

1.5 Thesis Organization

After the high-level discussion about the background and contributions of this thesis, the organization of the rest of the dissertation is as follows;

- Chapter 2 presents a broad view of human behavioral understanding. First, it defines human behaviors and existing approaches to model human behaviors. Second, it provides an overview of the state-of-the-art technologies typically used for sensory data acquisition to model different types of human behaviors. Third, it gives an overview of occupancy detection techniques. Also, it discusses the processing as well as the semantic enrichment of spatio-temporal data to study human behaviors. Lastly, existing applications of spatio-temporal data for risky movement detection and identification of intrusions for worker safety management are presented. The content of this chapter is partially extracted from the following publication;

  Arslan, M., Cruz, C. and Ginhac, D., 2019. Semantic enrichment of spatio-temporal trajectories for worker safety on construction sites. Personal and Ubiquitous Computing, p.1-16.

- Chapter 3 discusses the proposed framework 'OBiDE' to study human behaviors in dynamic environments. To study human behaviors, the capturing of the occupants` movements and occupancy were chosen. Later, for analysing the occupant movements by using the proposed framework, BLE beacons were selected. The content of this chapter is extracted from the following publication;

  Arslan, M., Cruz, C. and Ginhac, D., 2019. Understanding occupant behaviors in dynamic environments using OBiDE framework. Building and Environment 166, p.106412.

- Chapter 4 first discusses the preliminary measurement data analysis of beacon data for achieving better accuracy in identifying the building locations using BLE beacons. Second, the processes including a) trajectory data collection using BLE beacons, b) pre-processing as well as c) semantic enrichment of user movements with related contextual data of buildings for targeted case-studies are described. In the end, a method is mentioned which was used to detect the errors in the BLE beacon data for capturing occupant movements. The content of this chapter is extracted from the following publications;

  1) *Arslan, M., Cruz, C., Roxin, A.M. and Ginhac, D., 2018. Spatio-temporal analysis of trajectories for safer construction sites. Smart and Sustainable Built Environment, 7(1), pp.80-100.*
  2) *Arslan, M., Cruz, C. and Ginhac, D., 2019. Visualizing intrusions in dynamic building environments for worker safety. Safety Science, 120, pp.428-446.*
  3) *Arslan, M., Cruz, C. and Ginhac, D., 2019. Spatio-temporal dataset of building occupants. Data in Brief, p.104598.*

- Chapter 5 is based on the 1st case-study to use the OBiDE framework to analyse worker movements for construction safety management. The developed system 'WoTAS' uses the semantically-enriched worker trajectories and executes Hidden Markov Model to segment the worker movements into different movement states. The movement states were defined using different values of step length and turning angle. Later, the unsafe worker movements were visualized using BIM. A brief discussion is presented on how the developed prototype system functions with the aim of reducing the fatalities on construction sites. The content of this chapter is extracted from the following publication;

  Arslan, M., Cruz, C. and Ginhac, D., 2019. Semantic trajectory insights for worker safety in dynamic environments. Automation in Construction, 106, p.102854.

- Chapter 6 is based on the 2nd case-study to use the OBiDE framework to identify near-miss incidents such as intrusions on the construction sites. First, it models the required worker behaviors using the OBiDE framework. Second, a prototype system 'VIDEWS' is built to detect intrusions on construction sites for safety management. Lastly, BIM is used to generate the visualizations for detecting the locations on a building model where the intrusions have occurred. The content of this chapter is extracted from the following publication;

  Arslan, M., Cruz, C. and Ginhac, D., 2019. Visualizing intrusions in dynamic building environments for worker safety. Safety Science, 120, pp.428-446.

- Chapter 7 is the last part of this thesis which covers the extract of the conclusions made in this research study. Also, cost-benefit analysis is also performed to give some rough estimates of system development along with their perceived benefits. Moreover, future works are also proposed which can be used to further exploit the proposed OBiDE framework for an enhanced understanding of the human behaviors for different construction and facility management applications.

Chapter 2 - Understanding human behaviors

Outline

This chapter presents an overview of human behavioral understanding. First, it defines the human behaviors and the existing approaches to model occupant behaviors. Second, it provides an overview of the state-of-the-art technologies typically used for sensory data acquisition to model different types of human behaviors. Third, it gives an overview of occupancy detection techniques and discusses the processing as well as the semantic enrichment of spatio-temporal data to study occupant behaviors. Finally, it describes a few applications of spatio-temporal data to study worker behaviors on construction sites for safety management.

2.1 What are Occupant Behaviors?

Behaviors are observable actions or reactions of a user in response to external or internal stimuli (Wagner et al., 2018; Yan et al., 2015; Chen et al., 2015). These actions or reactions can be categorized into four main types (see Fig. 2.1) which are; *physiological adjustments* (e.g. sweating, shivering, etc.), *individual adjustments* (e.g. selection of clothes, using earplugs, etc.), *environmental adjustments* (HVAC adjustment, window opening or closing, etc.) and *spatial adjustments* (moving from one building facility to another, etc.) (Chen et al., 2015).

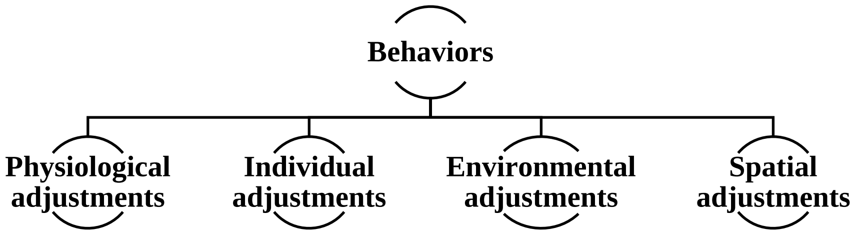

Fig. 2-1 Categorization of different behaviors

In other words, behaviors are the interactions (leaving or entering a room, visual and thermal indoor conditions adjustment using windows or blinds, doors, etc.) of building occupants which can be categorized into different movements, simple presence or actions with their environment (building, its systems and appliances) which impact on the building performance (heating or cooling, indoor air quality, energy, comfort, etc.) during their stay in a building (Wagner et al., 2018). Thereby, an occupant interaction which results in changing a building state (presence or absence in case of occupancy monitoring) or no interaction leaving the present state of a building unchanged are both facets of occupant behaviors (Wagner et al., 2018; Yan et al., 2015; Chen et al., 2015). Existing literature contains numerous systems to model occupant behaviors shown in Table 2.1.

Table 2-1 Examples of different types of occupant behaviors

Use case	Year	Type of behavior	Modeling approach used	Environment (Simulation or an actual site)	Sensor data, technologies or methods used
Behavioral modeling for building performance simulation (Cecconi et al., 2017)	2017	Occupancy	Probabilistic	University	Air temperature, global horizontal solar radiation and occupancy data
Modeling window control behavior	2017	Window control	Probabilistic	Residential apartment	Temperature, humidity, illuminance, CO_2

(Barthelmes et al., 2017)					concentration, and wind speed
Modeling building emergency evacuation plans (Rozo et al., 2019)	2019	Movements	Agent-based	University	n/a
Understanding the effect of worker-management interactions (Zhang et al., 2019)	2019	Worker's shortcut-taking behaviors	Agent-based	Simulation	Multi-agent modeling software
A proactive workers' safety risk evaluation framework (Chen et al., 2019)	2019	Position and posture	-	Building (Indoor)	Inertial measurement unit (IMU) and 3D skeleton data recording
A supervised learning approach for determining factors for unsafe behaviors (Goh et al., 2018)	2018	Working at heights	Machine learning-based	Tunnel construction project	Data collected based on the theory of reasoned action (TRA)
Safety harness detection (Fang et al., 2018)	2018	Falls from heights	Machine learning-based	Construction site	Convolutional neural networks, images and video recordings
Extracting and classifying unsafe behaviors linked with climbing a ladder (Ding et al., 2018)	2018	Falls from heights	Deep hybrid learning-based	Building (Indoor)	Convolution neural networks, long short-term memory, images and video recordings
Monitoring construction workers' vigilance (Wang et al., 2019)	2019	Obstacle avoidance	Structured surveys	Construction site	Wavelet packet decomposition, hybrid kinematic-EEG data
Detecting workers under varying poses against changing backgrounds (Son et al., 2019)	2019	Worker postures	Machine learning-based	Construction site	Convolutional neural networks, images from a movable digital camera
Tracking workers' gaze positions (Jeelani et al., 2018)	2018	Eye movements of workers	Computer vision-based	Construction site	Eye-tracking glasses, images and video recordings
Skeleton-based approach for identifying unsafe	2018	Worker body postures	-	Laboratory	Kinect 2.0 camera, video recordings

behaviors (Guo et al., 2018)					
Measuring workers affinity for or aversion to risky behaviors near hazardous zones (Rashid et al., 2018)	2018	Risk-taking behaviors	Machine learning-based	Open field	Smartphone GPS, random walks data
Identifying workers' near-miss falls (Zhang et al., 2018)	2018	Falls from heights	Artificial neural network (ANN)-based	Open field (test platform)	Smartphones with triaxial accelerometers, gyroscopes, three-axis acceleration and angle data
A robotic wearable exoskeleton for worker safety (Cho et al., 2018)	2018	Worker body postures		Building (Indoor)	Inertial measurement unit (IMU) system, motion data
Real-time locating for non-hard-hat workers (Zhang et al., 2019)	2019	Workers without safety hats	Internet of Things (IoT)-based	Laboratory and a real construction site	RFID triggers, thermal infrared sensor data
Monitoring harness use in construction (Gomez-de-Gabriel et al., 2019)	2019	Working at heights	Internet of Things (IoT)-based	Construction site	Bluetooth Low Energy (BLE) beacon and laser imaging detection and ranging (LIDAR) sensor data
Planning labor evacuation for construction sites (Marzouk et al., 2018)	2018	Movements	Agent-based	Residential building	Building Information Modeling (BIM)
Simulating the motions of heterogeneous human agents in case of emergency (Poulos et al., 2018)	2018	Movements	Agent-based	School building	Digital video cameras
Graph-based building simulation and data assimilation (Rai et al., 2018)	2018	Occupancy	Agent-based	Airport terminal	Video cameras
Predicting people's presence in buildings (Mahdavi et al., 2015)	2015	Occupancy	Non-Probabilistic	University	Motion detectors
Predictability of occupant presence	2017	Occupancy	Probabilistic	Laboratory	Webcams

using random walk modeling (Ahn et al., 2017)					
Predicting occupancy diversity factors (Davis et al., 2010)	2010	Occupancy	Non-Probabilistic	University	Security cameras and doorway counting sensors
Estimating the occupant activity level (Wolf et al., 2019)	2019	Activity	Probabilistic	Summerhouse and school	CO_2 sensors
Window opening model using deep learning methods (Markovic et al., 2018)	2018	Window control	Probabilistic	Office building	Indoor climate and air quality monitoring sensors

The studies presented in the above table clearly shows that for modeling the occupant behaviors and their interactions with the building, the occupants` movements and presence are the preconditions for any kind of behavior understanding as building occupants can only interact with the building environment if they are present inside the building. These studies are going to be discussed in the following sections.

2.2 Occupant Behavior Modeling

Primarily, there exist four major types of approaches to model occupant behaviors which are: static-deterministic, static-stochastic, dynamic-deterministic and dynamic-stochastic (Wagner et al., 2018; Yan et al., 2015; Chen et al., 2015; Hong et al., 2015). The static models cannot capture the influences that a building and its occupants can have on each other (Yan et al., 2015). These models are designed for understanding non-adaptive behaviors e.g. turning off the lights when occupants are going for a holiday, etc. Conversely, dynamic models deal with two-way interactions between a building and its occupants and are suitable for the adaptive nature of occupants' behaviors e.g. turning on the lights, changing the heating or cooling of a building, etc. Deterministic models produce the same outcomes whenever a simulation is run and give the homogeneous and deterministic results. Whereas, stochastic models produce different output whenever a simulation is run because the modeling parameters are selected randomly (Yan et al., 2015).

Among all the approaches discussed above, the modeling approach which has been used extensively in the industry is static-deterministic modeling (Wagner et al., 2018; Yan et al., 2015). However, this modeling approach is not suitable for constructing a robust building design as the uncertainty of occupants` behaviors is not considered. To include the uncertainty of the building occupants` behaviors, stochastic models are recommended (Yan et al., 2015). The three types of most commonly used stochastic models are; (1) Markov chain models, (2) Bernoulli models and (3) survival models (Yan et al., 2015). Discrete-time and discrete-event are the two main types of Markov chain models that take into account the environmental conditions for predicting occupants` actions in the latest timestamp or an event (Yan et al., 2015). The major limitation of Markov chain models is that they are not feasible to apply to the entire population of occupants as the computation and modeling effort increases linearly as the number of occupants in a building increase. In contrast to Markov chain models, Bernoulli models are the most simplified memoryless stochastic models in which the probabilities of events are independent of the previous events (Yan et al., 2015). Hence, Bernoulli processes do not require much information for modeling occupants` behaviors. Bernoulli modeling is used for energy modeling at a large scale as its scope can be efficiently applied to the entire building (Yan et al., 2015). However, Bernoulli processes do not output individual occupant behavior and are not capable of predicting the timings of the individual occupant`s behaviors. The third type of modeling approach i.e. survival modeling is used for estimating the time duration until an event occurs in a building. For example, these models are used for estimating how long a building probably remains unchanged by its occupants (Yan et al., 2015). In addition to three basic types of modeling approaches as discussed above, there exists an extension of the Markov chain models which use agent-based modeling. Agent-based models predict the influence of occupants by modeling individuals, their mutual interactions and how they interact with their building environment (Yan et al., 2015). In agent-based modeling, a huge amount of information (defining roles and relationships between the agents) is typically required and thus increases the modeling complexity. The term complexity is defined as the number of details required for modeling which is dependent on size (number of model components) and a resolution (number of model variables) (Wagner et al., 2018; Yan et al., 2015). An agent`s description generally consists of their attributes, resources, behavioral rules, etc. An extensive range of human agents present in the literature which include agents subject to reinforcement or belief-based learning, non-adaptive agents, and agents with capabilities of evolving new behaviors (Wagner et al., 2018; Yan et al., 2015).

2.3 Data Acquisition for Understanding Behaviors

This section presents methods of sensor data acquisition, different types of sensors used to acquire behaviors, considerations for choosing the appropriate sensory technology, and existing occupancy detection techniques.

2.3.1 Methods to acquire data

After selecting the most appropriate behavior modeling approach as per the application requirements, to create an understanding of the occupant interactions, the modeling process conventionally initiates from the sensor data acquisition of occupants (Wagner et al., 2018; Yan et al., 2015) along with the building environmental or infrastructural parameters. The methods for collecting occupant behaviors can be divided into three main categories, which are; 1) physical acquisition which can be a) on-site data collection or laboratory studies in controlled environments, 2) surveys, interviews and focus groups, and 3) virtual reality experiments (Yan et al., 2015). Physical acquisition studies involve monitoring building occupants in their physical environment. This environment can be an actual site where the occupants are present or a pre-fabricated environment in a laboratory setting that is controlled for a specific time to investigate the occupant behaviors. However, laboratory studies are expensive to build and may infer the behaviors differently as in real buildings the stress level of the occupants is higher (Yan et al., 2015). Moreover, the visibility of monitoring sensors installed in laboratories makes occupants conscious that they are been monitored which ultimately constrains their behaviors (Yan et al., 2015). Surveys and focus groups depend on the self-reporting of individual behaviors by filling out the questionnaires or through interviews. This cost-effective method enables the acquisition of behaviors that are not capturable using sensors (for instance; perception, attitudes, etc.). However, existing studies (Yan et al., 2015) show that the reported behaviors of building occupants may not always correlate with their actual behaviors. In addition, misunderstanding of the questions in the surveys may also lead to incorrect reporting of the information by the occupants (Yan et al., 2015). Apart from these two traditional approaches, the most emerging approach for occupant behavior modeling is virtual reality-based environments. This approach provides greater control of the environment in terms of environmental conditions (e.g. building layouts). However, there is limited support for visual and air quality configurations (Wagner et al., 2018; Yan et al., 2015).

2.3.2 Different types of sensors to acquire behaviors

An extensive range of different types of sensors (wired and wireless) for monitoring occupants and the environment are present in the literature (Wagner et al., 2018; Yan et al., 2015; Chen et al., 2015; Hong et al., 2015) to acquire information for modeling occupant behaviors and their interactions (energy consumption, etc.) with the buildings. Some of them are mentioned below;

- Mechanical sensing such as door badges, piezoelectric mats, etc.
- Image-based sensing using visual information captured from the cameras.
- Motion-based sensing such as passive infrared (PIR), ultrasonic Doppler, etc.
- Manual observations need humans for collecting datasets
- Wireless RF sensing using ambient sensors for temperature, humidity, light, etc.
- Consumption sensing for measuring water, electricity and gas usage in buildings

2.3.3 Considerations for the sensor technology adoption

For selecting a sensor technology for data acquisition, there exist nine factors (Wagner et al., 2018; Yan et al., 2015) which needs to be considered which are; 1. cost (including acquiring, installing and operating costs of the sensors), 2. power type (i.e. battery or power supply), 3. accuracy (i.e. difference between sensed data values and the ground truth), 4. sensor coverage range (i.e. the distance and the view angle covered by the sensors), 5. data collection type (i.e. event-based or periodic), 6. data storage (i.e. onboard or external storage of sensors), 7. deployment region (i.e. building inside or outside), 8. deployment level (i.e. floors, rooms, etc.) and 9. data sensed (i.e. binary data or value-based sensing). The quality of data captured from various types of sensors differ greatly in terms of the resolution of the deployed sensors (see Fig. 2.2). The spatial, temporal and occupant resolutions are combined for determining the overall resolution of the system for capturing the occupant behaviors (Wagner et al., 2018; Yan et al., 2015). The spatial resolution is defined in terms of building infrastructure (i.e. floor, rooms). Whereas occupant resolution can have 4 different levels of information which are; 1) simple detection based on the occupant presence in 0 or 1 values, 2) counting the number of occupants, 3) identifying the occupants, and 4) recognizing the occupants` activities. Along with spatial and occupant resolution, temporal resolution defines the smallest time period in which variations in spatial and occupant resolutions can be informed by a deployed sensor. As the resolution of the captured sensor data increases, the building space gets smaller, the occupants become more distinct based on their identities and the information from the sensor data will be accessed faster (Wagner et al., 2018). For example, a low-resolution system will only capture the binary information (presence or absence) of the occupants in a specific time where the identities of the occupants are not recognized. Whereas, a high-resolution system will be able to detect the number of occupants, their identifications, as well as their activities.

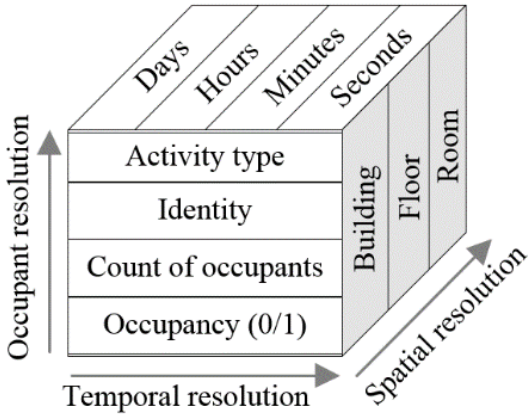

Fig. 2-2 Occupancy resolution (Melfi et al., 2011)

2.3.4 Occupancy detection techniques

While the previous section provides a brief overview of different types of sensors available for capturing different occupant behaviors, but occupants` movements and presence are the prerequisites for any type of occupant behaviors` understanding. Existing studies on occupancy detection and proximity analysis have been dominated by a wide range of sensors including passive infrared, Radio-Frequency Identification (RFID), ultrasonic, acoustic recognition, image cameras, Wi-Fi-based, Global Positioning System (GPS) and Bluetooth Low Energy (BLE) technology (Huang et al., 2019). Each technology has its own advantages and disadvantages which can be found in Huang et al. (2019) research. The output of a passive infrared sensor is binary and used for detecting the occupant presence instead of calculating the precise number of occupants in a facility (Mikkilineni et al., 2019). RFID-based sensors are conventionally used for calculating coarse-grained locations and occupancy (Lam et al., 2019). Though, as each RFID tag is mapped with an occupant, the privacy of an occupant is the main concern while using this technology (Huang et al., 2019). Occupancy detection using ultrasonic technique presents several disadvantages such as complexity in sensor configuration and management (Ghosh et al., 2019). Audio-based occupancy monitoring is inexpensive as the main equipment that is required consists of only microphones with microcontrollers. However, acoustic-based sensors are seldom used for occupancy monitoring as non-human sourced sound waves originating inside the noisy buildings can generate many errors in the collected occupancy data (Huang et al., 2019). Image cameras are also used for monitoring the building occupancy and estimating their locations in a building (Feng et al., 2019). However, the restrictions in the line-of-sight image collection introduce an increased complexity in deploying the cameras inside the building to cover all the possible room locations. Moreover, the expensive hardware for cameras and the issues related to occupant privacy have critically degraded its acceptance and deployment (Huang et al., 2019; Feng et al., 2019). In recent years, Wi-Fi-based proximity sensors are used extensively for calculating the indoor occupancy number (Shtar et al., 2019; Ali et al., 2019). While using this method, the Wi-Fi network is mandatory in all locations inside a building. In addition, each occupant needs to carry a Wi-Fi-enabled device all the time. If the occupants leave their devices in the offices and go to some other places. The Wi-Fi-based sensors will calculate the occupants` positions of their devices but not their actual locations (Huang et al., 2019). The errors in calculating the precise occupancy and location detection can arise in this case. Most smart devices include a GPS antenna that is used for locating the positions of their owners. However, the performance of GPS-enabled devices degrades severely inside the buildings because of the signal interferences caused by the building objects (Gong et al., 2019). As a result, Bluetooth-enabled beacons came into the application for capturing the movements of users for indoor building environments and have been used widely in recent studies (Paek et al., 2016).

Table 2-2 Comparison of location data acquisition technologies (Gomez-de-Gabriel et al., 2019)

	BLE Beacons	GPS	ZigBee	WiFi
Power usage	↑	↑↑	↑↑	↑↑↑
Coverage	↑↑	↑↑↑↑	↑↑↑	↑↑↑
Primary advantage	Low battery-powered	Long-range (outdoor)	Low price	Commonly used devices
Deployability	Easy	Easy	Easy	Medium
Deployment cost	↑	↑↑↑	↑↑	↑↑↑

↑↑↑↑ Very High -↑↑↑ High - ↑↑ Medium - ↑ Low

2.4 Spatio-temporal Trajectories and Semantic Enrichment

A few of location detection techniques to acquire spatio-temporal data for occupancy monitoring are mentioned in Table 2.2. The data collected from a typical location acquisition technology consists of spatio-temporal points having position information (latitude, longitude coordinates) with timestamps (Zheng, 2015). Based on the application requirements, these raw locations data points are transformed into a finite number of meaningful episodes called trajectories (Zheng, 2015). Formally, a trajectory is a discrete-time approximation extracted from a raw continuous movement track of a moving object having its Begin-End (t_{begin}, t_{end}) time interval (Zheng, 2015; Yan et al., 2013). Using the output (i.e. location dataset) of a geo-positioning system, $G = \{p_1, p_2, p_3 \dots, p_n\}$, where each point (p_i) corresponds to a unique spatio-temporal tuple in the form of $(longitude, latitude, timestamp)$. The sampling of the location data is achieved in a way that there should be no significant spatio-temporal gap between the tuples (Zheng, 2015; Yan et al., 2013). The information in such trajectory segments can only be used in locating some moving objects (e.g. what is the location of a worker1 at 10 am?) or performing statistical computations using spatio-temporal features of raw trajectories (e.g. which trajectories show an average speed over 1.4 meters per second?) (Zheng, 2015; Yan et al., 2013). However, most information systems for different applications require additional information from the application context. For example, understanding the movements of workers in a building needs additional information about the building environment (e.g. POIs such as rooms, etc.). For enriching trajectories with the related contextual information using external data sources both openly available and private data related to the buildings and the construction sites, a process of semantic enrichment is required (Yan et al., 2011). A brief comparison of existing trajectory models and their applications are mentioned in Table 2.3.

In the existing literature, there exist three major areas (Albanna et al., 2015; Zheng, 2015) on semantic trajectories: trajectory construction, trajectory segmentation, and trajectory annotation (Yan et al., 2011) using semantic data sources. However, the focus of this research will be on the latter two areas. In general, there are two modes to construct a trajectory (Albanna et al., 2015; Zheng, 2015): (a) online mode, where trajectories are constructed in real-time, and (b) offline mode, where all trajectories` construction processes are done in an offline mode. Although, the literature on an online construction of trajectories is limited, there are many offline trajectory construction methods present in the literature. In these methods, location data is collected in advance.

Table 2-3 Existing trajectory models and their applications

Model	**Type of model**	**Building environment**	**Application***									
			A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**
MADS (Spaccapietra et al., 2008)	Design pattern and dedicated data type	Outdoor				X						
SeMiTri (Yan et al., 2011)	Object-relational	Outdoor	X	X		X	X			X		X
The Baquara (Fileto et al., 2013)	Ontological	Outdoor	X			X	X			X	X	X
CONSTAnT (Bogorny et al., 2014)	Relational	Outdoor	X			X	X			X	X	X
The Baquara[2] (Fileto et al., 2015)	Ontological	Outdoor	X	X	X	X	X			X	X	X
SemMobi (Wu et al., 2015)	Document-oriented	Outdoor	X	X		X		X	X			
FrameSTEP (Nogueira, 2017)	Ontological	Outdoor	X	X		X	X				X	X
SMOPAT (Semantic MObility PATterns) (Wan et al., 2018)	-	Outdoor	X	X		X						

***A**: Identification of locations; **B**: Clustering of trajectories; **C**: Trajectory reduction; **D**: Trajectory segmentation; **E**: Transportation mode detection; **F**: Prediction; **G**: Recommendation; **H**: Activity recognition; **I**: Mapping with the environmental information; **J**: Behavior categorization;

2.4 Spatio-temporal Trajectories and Semantic Enrichment

Once location data is collected, it will undergo various processing stages such as data cleaning, map matching and compression (Albanna et al., 2015; Zheng, 2015). However, these methods are not suitable for real-life applications, where movements of objects are continuously updating. To address the requirement of an online trajectory construction for real-life applications, a real-time solution is known as SeTraStream (Yan et al., 2011) is present having an ability to process raw trajectory data within a controlled time window and generate trajectories with start and end time in an online mode. Once trajectories are created, a process of segmentation is applied for dividing these trajectories into a set of episodes based on predefined criteria. The very first data model proposed by the authors (Spaccapietra et al., 2008), in which the segmentation process is used to divide a trajectory into a set of moves and stops. A stop is defined as a place where a moving object has spent some specific time. However, other than a time threshold, segmentation can also be achieved by other attributes such as velocity, acceleration, direction, density, and geographic artifacts (Buchin et al., 2010). Similarly, there exists an extended segmentation framework to segment trajectories based on the movement states (Dabiri et al., 2018; Balzano et al., 2014). Their framework depends on the mapping of each movement state with relevant spatio-temporal criteria based on the expert knowledge, and manual user input. Moreover, Sankararaman et al. (2013) presented an approach to distinguish between similar and dissimilar portions in trajectories. After then, trajectories are divided into segments to extract contiguous portions of trajectories which are shared by many of the other trajectories. Furthermore, segmentation can be done based on representativeness (Panagiotakis et al., 2011). Such techniques perform global voting algorithms based on the local density and extract most representative sub-trajectories.

Once the segmentation of trajectories is completed, annotation techniques are applied to transform location trajectories into semantic trajectories (Albanna et al., 2015; Zheng, 2015). The annotation process involves enrichment of trajectory episodes with the meaningful information such as; the mapping of the POIs that can be in the form of points, in the form of lines or the geographical regions (Yan et al., 2011). Many annotation approaches are present in the existing literature as provided in Table 2.4. Wu et al. (2015) have used historic social media data to map it with user trajectories to understand the purpose of the trip. Based on the location history, relevant words are extracted from the Twitter data according to the mobility records, and user interests are retrieved for visiting a specific location at a specific time. In addition, user activities have also been used to annotate raw trajectories (Furletti et al., 2013). However, to cover a larger pool of the POIs, there is a need for integration with more datasets to enable tracking in large cities for the extraction of user activities (Furletti et al., 2013).

Nogueira et al. (2016) have developed a framework to annotate trajectories based on episodes. It basically defines the environment where the user trajectory has been taken placed based on the Linked Open Data (LOD) cloud and OSM data. The ability of this framework of describing the spatial context of GPS trajectories can be used as a building block of future expert systems for trajectory exploration. Moreover, trajectories of slow and fast-moving objects are also annotated at different levels of data abstraction using a multi-layer framework (Yan et al., 2011). The locations where objects move provide information about their interests. At the same time, such behavioral analysis gives information about the popularity of visited places. A similar research effort is presented by Graaff et al. (2016), where an algorithm is proposed for combining existing multiple trajectory pre-processing methods to identify the visited POIs for detecting different indoor activities in an urban setting. Furthermore, the ontological modeling approach has also been used to abstract trajectory data in a multilevel hierarchy using LOD collections and social media data (Fileto et al., 2015). This integration can enable to query trajectories for mobility analysis based on the domain and application-related knowledge. In addition to automatic semantic annotations, there also exist dynamic and clustering-based semantic annotation methods based on the contextual and geo-localization data in the literature (Wan et al., 2018; Wu et al., 2015; Cai et al., 2018). Such methods calculate the local density of words and map words to each trajectory record, hence providing visualizations for trajectory exploration.

Table 2-4 Comparison of existing semantic annotations of trajectories

Use case	Environment	Findings	Key methods used	Type of data	Annotation types comparison			
					Social media	Point	Region	Line
Semantic annotation of mobility data using social media (Wu et al., 2015)	Outdoor	Extracted purposes and interests of a user from his location history.	Kernel density estimation model	Geo-tagged tweets	Y	Y	N	N
Inferring human activities from GPS tracks (Furletti et al., 2013)	Outdoor	Automatically annotated trajectories based on user activities.	Gravity law	GPS trajectories of a car	N	Y	N	N
Annotating semantic trajectories based on episodes (Nogueira et al., 2018)	Outdoor	Environments are identified where trajectories took place.	Linked Open Data cloud and OSM	GPS trajectories of a jogger	N	N	Y	N
Semantic annotation of heterogeneous trajectories (Yan et al., 2011)	Outdoor	Annotated trajectories for any kind of moving objects.	Java 6 (PostgreSQL)	GPS records of taxis and private cars	N	Y	Y	Y
Automated semantic trajectory annotation with indoor POI visits	Outdoor	Combined multiple trajectory pre-processing techniques to extract POIs.	-	Trajectories formation using uploaded	N	Y	N	N

(de Graaff et al., 2016)				pictures				
Automated semantic annotations based on existing knowledge bases (Fileto et al., 2015)	Outdoor	Abstraction in a multilevel hierarchy of progressively detailed movement segments.	geoSPARQL and PostGIS	User's trail from Flickr and tweets	Y	Y	Y	Y
Dynamic semantic annotation of trajectories (Wu et al., 2015)	Outdoor	Annotation using contextual social media data.	Kernel density estimation model	Geo-tagged tweets	Y	N	N	N
Mining semantic trajectory mobility patterns (Wan et al., 2018)	Outdoor	Characterization of the semantic mobility of vehicles is achieved by tagging visit purpose to each trajectory.	Google Maps and prefix span algorithm	Private vehicle trajectories	N	N	Y	N
Finding semantic level trajectory behaviors through semantic trajectory clustering (Cai et al., 2018)	Outdoor	Extracted common semantic trajectories using an extended OPTICS algorithm.	Density-based clustering algorithm	Geo-tagged photos from Flickr	Y	N	Y	N

2.5 Spatio-temporal Trajectories for Worker Safety Management

The previous section gives a brief overview of spatio-temporal trajectories and different semantic enrichment techniques which are conventionally used in the state-of-the-art literature to enrich occupants` movements and presence with the building environment information to conduct advanced spatio-temporal behavioral analysis by inferring different occupant activities for various building management applications. Though, the focus of this research is to study site occupant (i.e. worker) behaviors in the construction environment for safety management. According to the OSHA Annual Statistics Report (BLS, 2017), a major number of fatalities in the U.S. construction industry were resulted from “unsafe movement-related” (falling from heights (i.e. 39.2%) and struck-by-object incidents (i.e. 8.2%)) worker behaviors. For example, a) limited spatial awareness of the operating construction machinery involving sharp movements and rotations within the workers` proximity due to blind spots and surrounding noise, and b) an unauthorized staying in or stepping into the hazardous areas (also called as intrusions) without realizing the potential dangers associated with the locations and are very common type of near-miss incidents as defined in the previous chapter. One of the most effective ways to reduce the occurrences of unsafe worker behaviors (risky movements and intrusions) for reducing the chances of near-misses which eventually lead to accidents is by monitoring the worker movements in real-time. The state-of-the-art literature is reviewed in Tables 2.5 and 2.6 to explore existing solutions to monitor real-time worker movements for safety management.

Table 2-5 Existing solutions to monitor risky movements of workers

Use cases	**Building environment**	**Dataset type**	**Findings**	**Key components and technologies**
Autonomous crane safety monitoring (Luo et al. 2014)	Outdoor	Simulated data of movements	Provided an understanding of the impact of data errors on safety monitoring system performance.	Cricket motes
Continuous localization of construction workers (Park and Brilakis, 2016)	Outdoor	Location data	Automatic detection and tracking of construction workers in video frames.	Video cameras
Tracking the real-time position of workers on construction sites (Li et al. 2015)	Indoor and outdoor	Location data	Achieved cost-effective real-time location system using chirp-spread-spectrum for construction safety.	NanoPAN 5375RF module for data transfer, Chirp spread spectrum (CSS) technology (IEEE 802.15.4.a)
Predicting movements of onsite workers and mobile equipment (Zhu et al. 2016)	Indoor	Location data	Achieved high predicition accuracy based on the worker or equipment's previous movements.	Kalman filters, video cameras
Real-time location tracking of multiple construction laborers (Lim et al. 2016)	Underground construction site	Acceleration and positioning data	Minimized the tracking error using accelerometer and BLE technology together.	Bluetooth Low Energy (BLE) and accelerometer
Real-time building protocol control and data visualization (Costin et al. 2015)	Indoor	Location data	Achieved RFID-BIM integration to produce leading indicators for monitoring safety, security, etc.	RFID and BIM
Representing spatiotemporal orders of construction processes (Yang et al. 2017)	Indoor	Location data	The identification of the interactions among site workers is achieved using the interdependence network.	Real-time location system (RTLS)
Modeling, visualizing and analyzing workspace	Outdoor	Location data	Identification and visualization of the required or potentially congested	Global Positioning System (GPS) and BIM

requirements (Zhang et al. 2015)			workspaces are achieved.	
Inferring slip, trip and fall (STF) safety hazards (Yang et al. 2019)	Indoor	Location, acceleration, orientation, and rotation data	Automatic identification of the STF hazards is achieved.	Ultra-wideband (UWB) location-tracking system
Detecting and visualizing dynamic workspaces of workers on foot (Luo et al. 2019)	Outdoor	Location data	Provided a solution to collect two types of data: 1) action classes, telling what workers are doing, and 2) action locations, indicating where they are.	Video cameras

Table 2-6 Existing solutions to monitor intrusions on sites

Use cases	**Building environment**	**Dataset**	**Findings**	**Key components and technologies**
Intrusion warning and assessment system (Heng et al., 2016)	Outdoor	Location trajectories of 3 workers for 6 hours in day	Developed a multi-user platform for obtaining the worker positions in relation to virtual hazardous zones.	Real-Time Locating System (RTLS) consisting of tags
Identifying and recording intrusion behaviors on construction sites (Shuang et al., 2019)	Outdoor	147 construction workers` data over a 4-month period	Analyzed the effect of age and gender of construction workers on the frequency of intrusions.	BIM and RTLS
Controlling interferences on sites using a real-time monitoring system (Naticchia et al., 2013)	Outdoor and indoor	600 positioning records sampled at 5 minutes.	Stored and maintained workers' site interactions and detected interferences.	IEEE 802.15.4 standard medium access
Preventing near-miss interactions between construction resources (Golovina et al., 2016)	Outdoor	70 min long trajectories of a worker and a skid steer loader.	Recorded, identified, and analyzed hazardous near-miss situations between workers and heavy construction equipment.	GPS
Detecting non-certified work on sites (Fang et al., 2018)	Outdoor	Public datasets (WIDER FACE dataset and CelebFaces+ dataset)	Non-certified workers are detected using video imaging technology and notifications are generated to cease their activities.	Vision-based deep learning algorithms

Preventing workers from accessing hazardous areas using a proactive system (Carbonari et al., 2011)	Indoor	-	Warnings are generated to inspectors before the occurrences of the interferences.	Ultra-Wide Band (UWB) technology
Constructing a model after analyzing historical accidents cases to prevent accidents (Wu et al., 2013)	Outdoor and indoor	499 accident cases from 1990 to 2008.	An integrated information management model is built to track stuck-by-falling-object accidents.	ZigBee and RFID technology
BIM and cloud-based indoor localization system (Fang et al., 2016)	Indoor	-	Performed real-time data processing and generated visualization for remote monitoring for detecting intrusions.	Passive RFID system
Implementing a proactive accident prevention solution (Yang et al., 2012)	Outdoor	4,640 accident cases from the U.S. OSHA database.	Analyzed the automatic identification requirements consisting of access control, training and inspection information and operation authority.	ZigBee and RFID technology
Presenting a safety early warning system in underground construction (Ding et al., 2013)	Underground	-	Designed and validated an IoT based safety system on the construction site to provide information about dangerous situations in advance.	A fiber Bragg grating (FBG) sensor system and a RFID
Proximity hazard indicator for near-miss interactions (Teizer and Cheng, 2015)	Outdoor	-	Automatically identified the areas of static and dynamic hazards on a construction site.	UWB localization system

As it can be observed from the Table 2.5 and 2.6, in the literature, to monitor worker behaviors in real-time for minimizing the near-misses, sensing and warning-based technologies have been utilized to overcome the limitations of BBS training. The developed systems based on sensing and warning-based technologies such as Radio Frequency Identification (RFID), location estimation using 802.11, Global Positioning Systems (GPS) and Indoor Positioning Systems (IPS) are used for;

1. acquiring spatio-temporal data of workers and the site environment to understand mobility dynamics of workers and operational machinery in relation to the site environment,

2. executing safety risk assessment methods on the acquired sensor data,
3. generating accurate and timely warnings to the construction workforce (construction workers and machinery operators) for triggering H&S interventions to improve worker behavior during hazard proximity conditions or unauthorized accesses, and
4. disseminating the information of near-misses to site supervisors and H&S managers for quick actions for maintaining site safety (Antwi-Afari et al., 2019).

The objective of the existing systems as stated above is not only to construct proactive surveillance methods for site supervisory staff and H&S professionals for protecting the H&S of frontline workers while working in hazardous situations on sites but also indirectly lowering the costs of construction projects by reducing the occurrences of accidents on sites (Teizer et al., 2010; Antwi-Afari et al., 2019).

2.6 Research Gaps Identified from the Literature Review

After reviewing existing systems to study unsafe worker behaviors, it is concluded that spatio-temporal data is extensively analyzed to avoid near-miss incidents on construction sites. However, previously designed systems do not offer a mechanism to recognize near-misses such as risky movements and intrusions in dynamic environments where the building locations change over time in terms of their position, size, properties, and relationships with the environment. New infrastructural support such as walls, etc. are often added often on construction sites, while others are detached (Arslan et al., 2019c). This opens more challenges to keep track of the changes in the contextual information associated with the locations which evolve for identifying intrusions from the perspective of the building or site environment. To fill this research gap, the most relevant modeling approaches for constructing semantic trajectories using spatio-temporal are reviewed. This literature review suggested that ontology-based modeling is the most preferred approach for enriching user trajectories with the relevant semantic information as ontologies provide the convenience of sharing, exchanging and reusing the domain-specific knowledge (Fensel, 2002). Moreover, after the literature review, a few research gaps have been realized. These gaps encompass;

1. Existing semantic trajectory enrichment models are designed specifically for outdoor environments for tagging the relevant semantic information with the moving persons or vehicles` trajectories. The extraction of the insights related to the behaviors of moving objects within the building settings is not yet adequately explored.

2. To the best of our knowledge, the data models present in the literature for performing semantic enrichments hold static information regarding the environment in which the objects are in motion. However, for modeling real-life trajectories and extracting real-time insights about the moving object behaviors updated contextual information related to the building environment is necessary by the trajectory data model to generate semantically-enriched trajectories.

3. The focus of the baseline models presented in Table 2.3 has been kept limited only to construct semantic trajectories. The feasibility of visualizing the semantic trajectories by integrating the trajectory data models` output with the existing open-sourced smart city solutions, for example, Building Information Modeling (BIM) for different industrial application scenarios is missing in the literature.

Chapter 3 - OBiDE Framework

Outline

This chapter discusses the need for constructing the framework to study occupant behaviors in dynamic environments. Currently, there exist numerous behavior modeling techniques in the literature for understanding the occupant behaviors. However, these techniques are application-specific, primarily aims to model energy-related occupant behaviors and do not provide the generalization of an occupant behavioral model which can be useable for a broad range of built environment applications. In addition, these models provide inadequate support to study occupant behaviors in dynamic environments. To construct a generalized framework for studying the occupant behaviors, occupants` movements and presence are identified as the pre-requisites from the literature review (see Chapter 2) for any kind of occupant behavioral understanding. To study the occupant movements and presence (i.e. the fundamental facets for any kind of occupant behavioral understanding) in a dynamic building environment context, the framework 'Occupant Behaviors in Dynamic Environments' (OBiDE) is proposed in this chapter. The OBiDE framework consists of six processes starting from; 1) defining human behaviors using DNAS ontology, 2) capturing spatio-temporal movements of building occupants, 3) preprocessing the acquired movements, 4) semantically enriching the spatio-temporal movements and historicizing them using the changing building environment context, 5) applying the behavior modeling techniques to achieve the desired classification of the occupant movements based on their inferred states, and finally 6) visualizing the classified occupant movements using a BIM software.

3.1 Occupants` Movements and Presence

After an extensive review of existing applications of occupant behavior modeling (see Chapter 2), it is observed that numerous systems have been developed to understand occupant behaviors. Each system is developed after conducting different types of measurements (real site and laboratory) and surveys that focus on incorporating different model variables (e.g. physical, biological and environmental, etc.) and human factors for representing different occupant behaviors (e.g. movements, occupancy, body postures, etc.) for different applications of facility management with respect to different types of buildings (residential, commercial, etc.) (Wagner et al., 2018; Yan et al., 2015; Chen et al., 2015; Hong et al., 2015). Consequently, the existing systems for understanding occupant behaviors cannot be compared to one another as each system has unique functionality and the scope of inferring the actions using the sensory data is limited to the application. However, Hong et al. (2015) provided a DNAS ontology (see Fig. 3.1), which acts as a technical framework to standardize the major components (Drivers, Needs, Actions, and Systems) required to model occupant behaviors. A DNAS ontology primarily aims to model energy-related occupant behaviors (Hong et al., 2015). However, the features which were observed fundamental in DNAS and most of the other developed occupant modeling systems (see Table 2.1) are occupants` movements and presence. The occupants` movements and presence are considered as the prerequisites for any kind of behavior understanding as building occupants can only interact with the building if they are present inside the building (Yan et al., 2015).

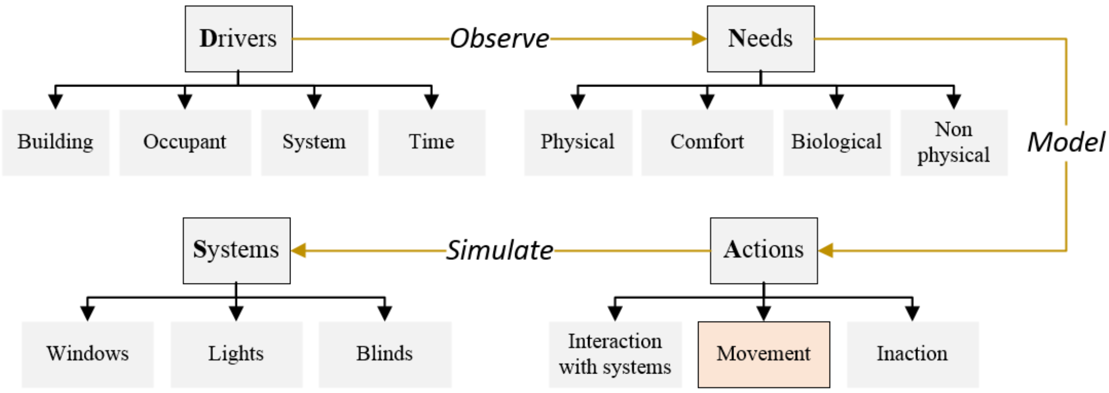

Fig. 3-1 A DNAS framework by Hong et al. (2015)

3.2 Occupants` Movements and Presence in Dynamic Environments

Existing systems in the literature are built by capturing the stochastic and reactive nature of occupant behaviors to model their movements and presence in buildings. Ultimately, these systems contribute to enhancing the understanding of occupant behaviors by increasing the occupancy resolution (i.e. inferring different occupant activities using their movements and presence) for different building monitoring and management applications (Melfi et al., 2011). However, the existing systems for occupant behavior modeling do not incorporate the information of complex dynamic environments where the building objects (occupants and building locations) evolve (see Fig. 3.2). Over time, the functionalities of the locations in a building often change (i.e. change in semantics or a context) (Arslan et al., 2019b). For example, a room named ‘inventory’ in a building is now an ‘office’ having different functionality. Likewise, due to the placement of certain inventory in the specific area of a building, the floor area of a building (a set of rooms and a corridor) became a ‘restricted area’. Also, new walls or infrastructure support may be added in a building (Arslan et al., 2019e). This will result in a change in the dimensions (i.e. geometry) of building locations (called as spatial changes). The change in the semantics of building locations often occurs in constructed facilities whereas, the spatial changes take place rarely. Such changes need to be incorporated in occupant behavior modeling as a change in the purpose or a position of building locations will result in different behaviors of occupants which ultimately represent different occupant activities (Arslan et al., 2019e). The updated spatial and semantic information about the building locations along with the previous information will contribute to an improved understanding of occupant behaviors concerning the changes occurred in the building environment. Resulted occupant behaviors after modeling the dynamicity of the building environment can be used for different construction and built environment applications such as; construction resource monitoring for improved safety, managing building spaces based on their utilization, implementing occupant access control systems, etc. To address these requirements of dynamic building environments that contain evolving building objects, an integrated framework named ‘OBIDE’ is proposed (see Fig. 3.3).

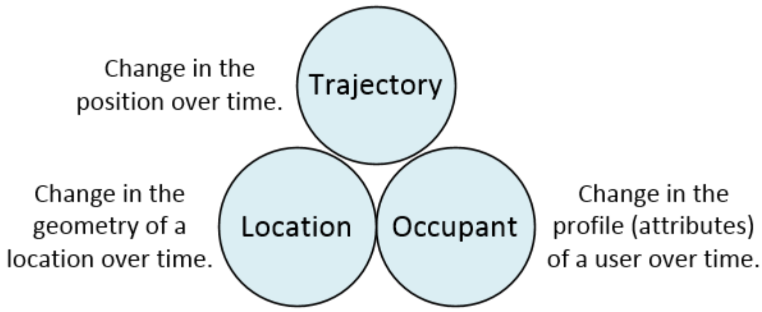

Fig. 3-2 Entities of a dynamic building environment

3.3 The OBIDE Framework

There are six different processes (see Fig. 3.3) which are involved in constructing the OBiDE framework (see Fig. 3.4) which are;

1) ***Defining human behaviors using DNAS ontology***; In DNAS, ‘Drivers’ indicate the environmental factors which contribute to stimulating occupants to achieve a physical, physiological or psychological need. ‘Needs’ define the requirements (physical or non-physical) that must be fulfilled to achieve desired satisfaction. ‘Actions’ are the set of interactions with the building systems to achieve the desired comfort, and ‘Systems’ represent the building and its equipment.

2) ***Capturing behaviors of occupants***; As occupants` movements and presence are the prerequisites for any kind of human behavior understanding, an OBiDE framework specifically deals with the capturing the spatio-temporal movements of building occupants in dynamic environments.

3) ***Preprocessing the behaviors***; Once the spatio-temporal movements of occupants are captured in the form of trajectories, these movements are filtered, segmented (i.e. based on the use case e.g. walk and run segments), and the building locations having the stay locations of occupants are identified (again, it is dependent on the use-case application).

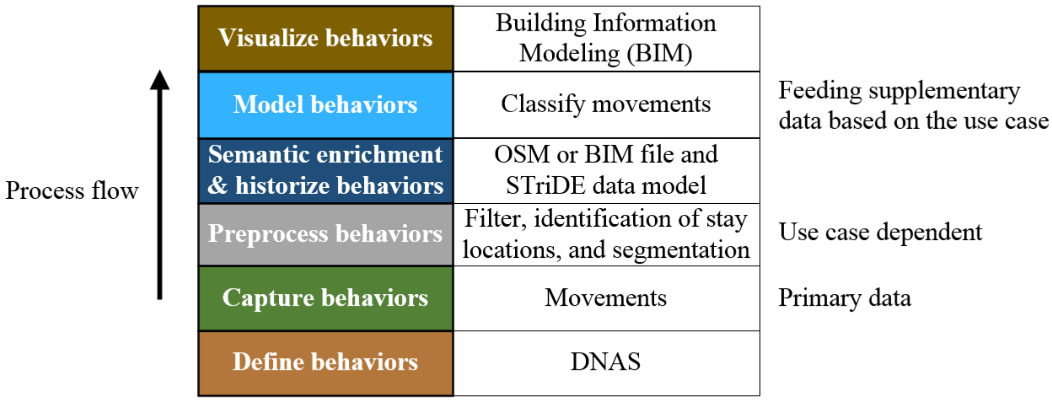

Fig. 3-3 OBiDE framework having different processes

4) ***Semantic enrichment and storing the movements***; After preprocessing the occupant movements, the movements are contextually enriched with the corresponding building locations to add more meaning to the collected trajectories using the building information extracted from BIM or an OSM file. Later, these spatio-temporal movements are stored in the STriDE data model.

5) ***Modeling behaviors***; Based on the use-case requirements, occupant behavior modeling techniques are applied to the processed spatio-temporal data along with the supplementary data of the occupants or the building environment to achieve the desired classification of the occupant behaviors. As discussed in the previous chapter, primarily, there exist four major types of approaches to model occupant behaviors which are: static-deterministic, static-stochastic, dynamic-deterministic and dynamic-stochastic (Wagner et al., 2018; Yan et al., 2015; Chen et al., 2015; Hong et al., 2015). Among all the approaches, the modeling approach which has been used extensively in the industry is static-deterministic modeling (Yan et al., 2015). However, this modeling approach is not suitable for constructing a robust building design as the uncertainty of occupants` behaviors is not considered. To include the uncertainty of the building occupants` behaviors, stochastic models are recommended (Yan et al., 2015). For an improved understanding of occupant behaviors by decoding and classifying their different types of movements, the stored occupant spatio-temporal data is later fed to a probabilistic model (e.g. Hidden Markov Model in our case) for inferring their movement states. As for recognizing and categorizing the movements, many case studies are present in the literature based on stochastic modeling (Arslan et al., 2019b). Though, the HMMs-based method is chosen for our study because statistical HMMs are applied widely in many works for categorizing the trajectory movements and extracting patterns. The computation of probabilities of different movements of occupants in terms of evolving building locations is used for enriching a DNAS ontology for an

enhanced understanding of occupant behaviors by categorizing their actions based on their movements for our two safety management applications (as discussed in Chapter 5 and 6).

6) ***Visualizing the classified human behaviors***; Based on the literature survey, BIM is chosen for visualizing the classified movements of occupants. The justification of using BIM is already described in Chapter 1.

The objective of the OBiDE (see Fig. 3.4) is to facilitate the development of new systems to enable the standardization of occupant behaviors` descriptions by incorporating the real-life dynamicity of building environments during simulations for an improved understanding of occupant behaviors (Arslan et al., 2019e). The proposed framework consists of an ontology-based data model named 'STriDE' which stores the movements of occupants as well as maintains the historicizations of building locations (Arslan et al., 2019d; Cruz et al., 2017). With the help of appropriate techniques (conventional or agent-based) to model the movements and presence of the occupants, the occupancy resolution of the model will be enhanced which ultimately helps to infer occupant activities with higher accuracy (Melfi et al., 2011). The STriDE model does not only store the spatio-temporal movements of occupants but also offers flexibility for further expansions and interoperability for the enrichment of occupant behaviors` descriptions with supplementary data sources to complement the process of occupant behavioral analysis.

3.3.1 Presentation of the STriDE model

The STriDE (see Fig. 3.5) models the building environment as a collection of different building objects (entities). In our case, we have three building entities (see Fig. 3.2) which are 1. a trajectory, which corresponds to an occupant location i.e. a set of spatio-temporal points, 2. a location (i.e. a physical building location) and 3. an occupant. Each building entity evolves under the action of different processes. The lifecycle of each building entity is summed up into a series of different states. Each state represents a change in the entity. As shown in Fig. 3.2, a change can occur in the location, geometry or the semantic (thematic) attributes of a building entity.

The different changes in building entities as mentioned above can occur independently or simultaneously in a dynamic building environment. The 'concepts' are defined in a data model for tagging the building locations with spatio-temporal trajectories of occupants. The data model (see Fig. 3.5) uses a set of classes and properties from the existing vocabularies for defining the concepts and their relationships which are; 1. Simple Knowledge Organization System (SKOS), 2. Dublin Core Terms (DCT) and 3. GeoSPARQL (GEO). Using these vocabularies, the concepts are defined and stored in the 'concept scheme'. A concept scheme is a collection of different concepts defined as a hierarchy that corresponds to different building locations (see Fig. 3.5). The purpose of defining as a hierarchy is to be more precise about the tagging of building locations to spatio-temporal trajectories of occupants. Using a set of different concepts, occupant profiles are created to define the access level of the occupants. For identifying the occupants, the tagged concepts with the occupant trajectory`s timeslices (TSs) are compared with the allocated concepts as per their profiles. To model the building environment, the STriDE model keeps track the building entities as well as different relations among them which are;

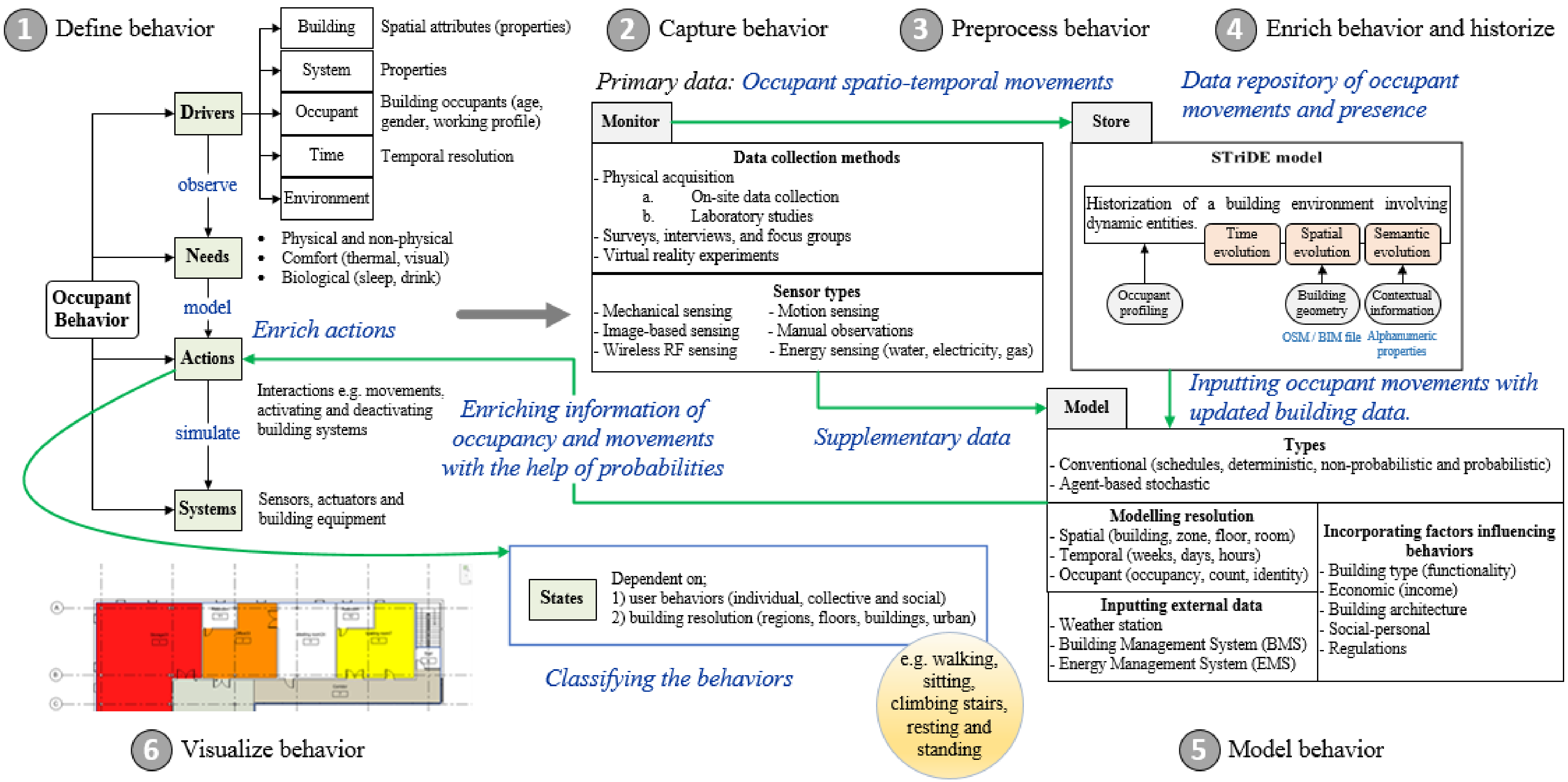

Fig. 3-4 Generalized OBiDE framework

1. ***the spatial relation*;** specifies how an entity is in a building in relation to a reference entity (a room). For example, an entity A (i.e. an object) passes through entity B (i.e. a location) or the geometries of entities A and B overlap each other or touch at a specific point in a building.
2. ***the spatio-temporal relation*;** specifies how building entities (two rooms, or a room and an occupant) are related to each other at the same time.
3. ***the filiation relation*;** specifies how building entities are related by ancestry or successor (Harbelot et al., 2015). It defines the succession links which exist between several representations of the same entity at different time instants. A data model deals with two types of filiation relationships; a continuation (an identity of a building entity remains the same while an entity changes) and a derivation (after a change a new building entity is created from the parent entity) (Harbelot et al., 2015). For more information, see Fig. 3.7, 3.8 and 3.10 below.

Timeslices for tracking the building evolution

For tracking the evolution of building entities (occupants, trajectories and rooms) and their relations with one another, the model uses the concept of TSs. A TS includes four components which are; an identity, alphanumeric properties (semantic component), a geographical (spatial) and a time component (Arslan et al., 2019e; Cruz et al., 2017). At the occurrence of a change in any component of a TS excluding the identity, a new TS is generated inheriting the components of the last known TS. To show how our model keeps tracking the evolution of dynamic building entities (location, trajectory, and occupant) using TSs, three possible scenarios are described which are;

- ***The functionality of a location is changed (a geographical component)*;** A building room named: 'Office' has a size of 24m^2 as shown in Fig. 3.6. The room is tagged as an 'Office' in a model having a defined geometry. After a while, a building room named: 'Office' is now a 'Meeting Room' having the same geometry of 24m^2 and tagged as a 'Meeting Room' with the help of a new TS Room1_1 in a model having the same geometry as of 'Office'. The transition between an 'Office' to a 'Meeting Room' is stored using a filiation link as shown in Fig. 3.7.

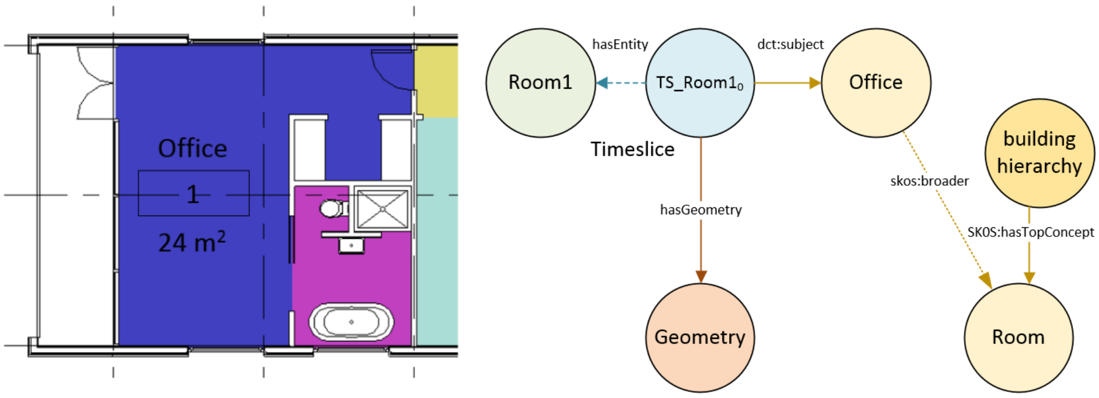

Fig. 3-5 A building room with its information

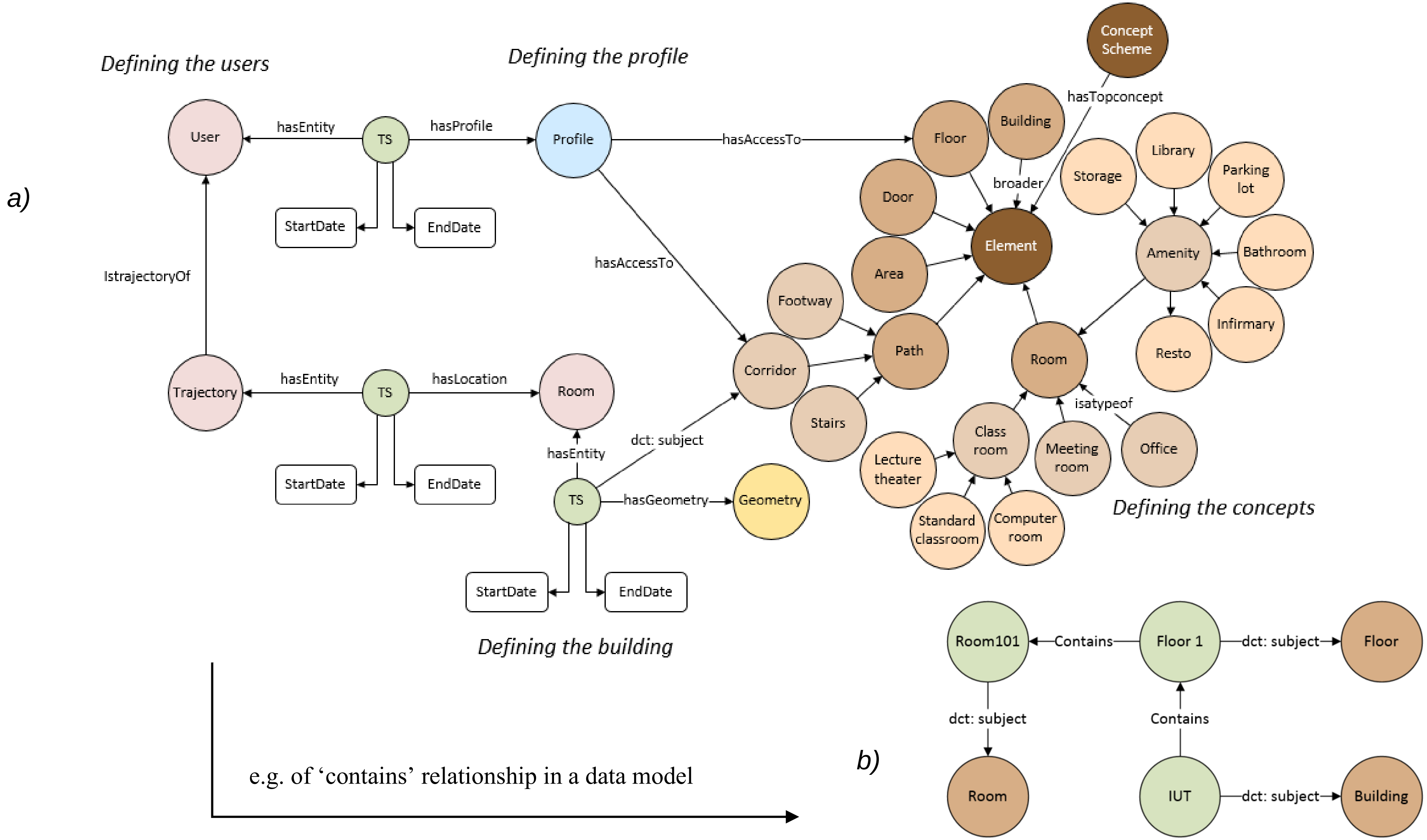

Fig. 3-6 (a) (top) The STriDE model (b) (bottom) An example of 'contains 'relationship (The "contains" relation is expressed at the instance level: The Institute of Technology (IUT) contains the 1st floor which contains the Room101).

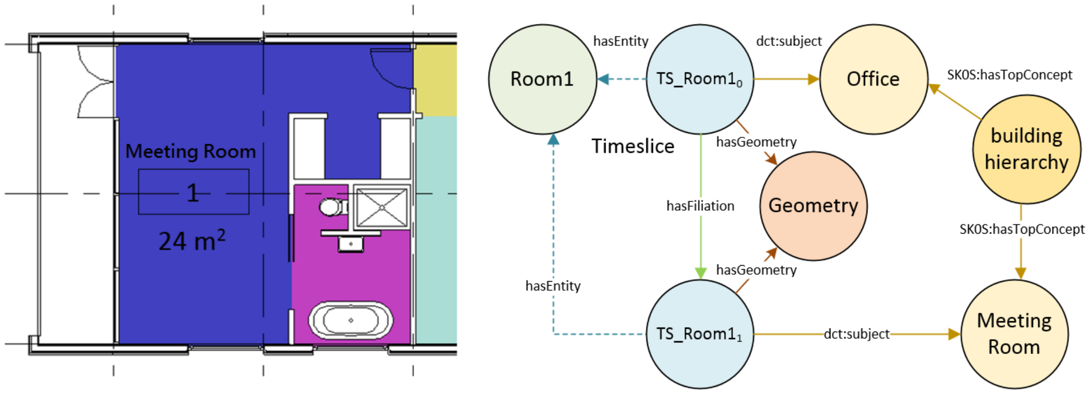

Fig. 3-7 The functionality of a building room is changed.

- ***The geometry of a building room is changed;*** Let's suppose, a building room named: 'Meeting Room' has now a size (geometry) of 28m^2. A new TS is created with the name of TS_Room1_1 to hold this change in geometry. The change in the geometry between an 'Office' to a 'Meeting Room' is stored using a filiation link as shown in Fig. 3.8.
-

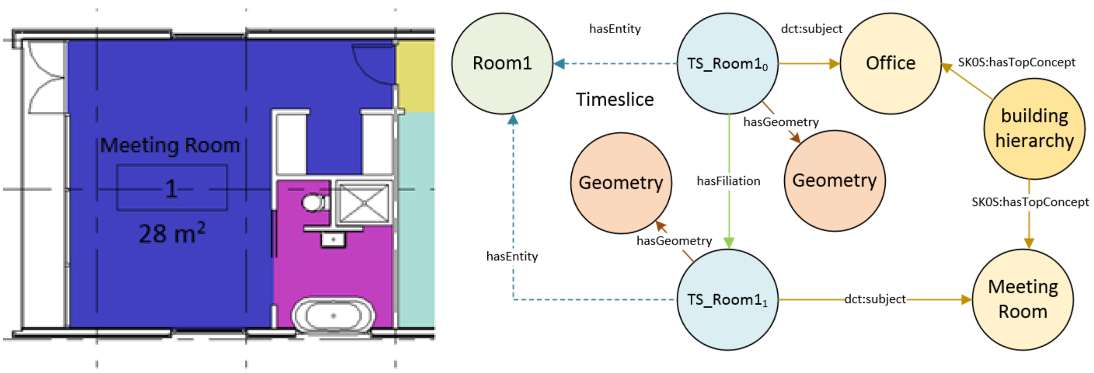

Fig. 3-8 The geometry of a building room is changed.

- ***The creation of a new building room;*** Let's suppose, a room labeled 'Office' is demolished. A new room i.e. 'Room2' is created as shown in a building model created in Revit Architecture software by Autodesk. A Revit software is used for creating the three-dimensional model of a building (Arslan et al., 2019b). A room in the Revit represents using a three-dimensional volume representing a real space. A Revit room can be created by going to 'room tools' located at the 'room and area' panel in the Revit software (see Fig. 3.9 (top row)). Later, this newly created room is labeled by placing the room tag associated with it as shown in Fig. 3.9 (bottom row). As a building room named: 'Office' is destroyed (see Fig. 3.10 (a and b)). Initially, its TS Room1_0 is updated by changing its end date-time stamp to show that a room is no longer exists (see Fig. 3.10c). Also, two new TSs are constructed which are Room2_0 and Room3_0 to show the construction of the two rooms having different geometries (see Fig. 3.10d). In this case, new rooms are created which are linked to the previous room using the filiation links.

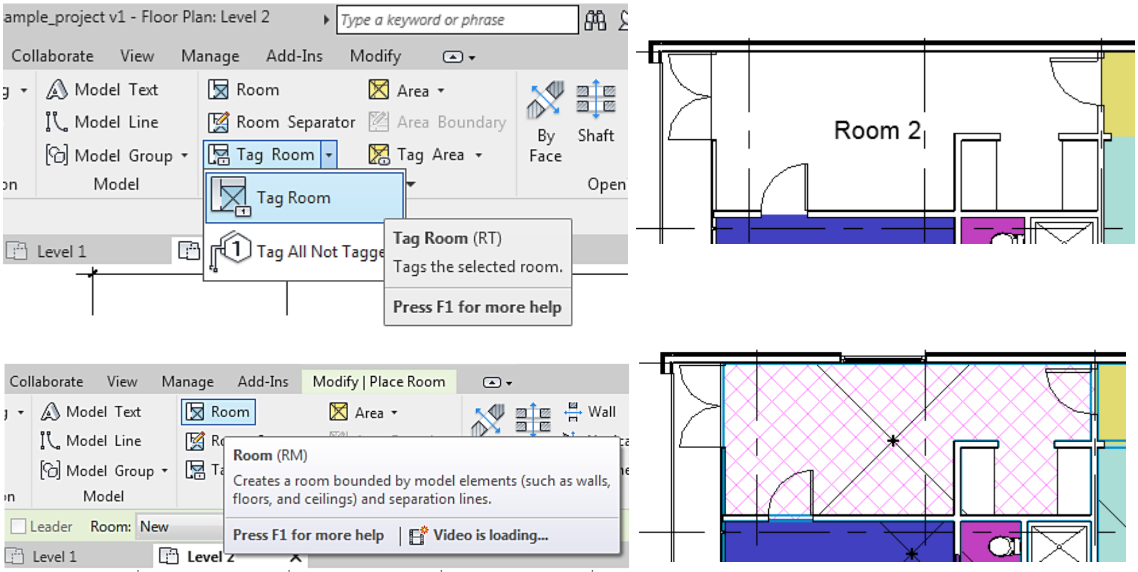

Fig. 3-9 Creating a room and tagging with a label 'Room 2' in a Revit software

3.4 Discussion

The OBiDE framework aims to add the dynamic context to occupant behaviors` modeling process as context (i.e. details about building space, time and environment) should be closely linked with the occupant movements for an enhanced understanding of their actions. After creating a knowledge base of historicized movements of occupants using the STriDE model by tracking the evolutions occurring in the building environment, an appropriate modeling technique as per application requirements (agent-based or conventional) can be applied to the movements for calculating the probabilities of different actions of occupants across different building locations. A critical question here is the degree of detail about occupants and the environment which should be included in the modeling stage to attain the targeted understanding of occupant behaviors. These details include occupant profiles, the type and the number of buildings, and the required temporal (e.g. hours, minutes, seconds) and spatial resolutions (e.g. floors, rooms). The STriDE model can execute an access control system after creating different occupant profiles with the help of concepts as described above. Also, the STriDE model can hold data of multiple buildings with the help of OpenStreetMap (OSM) building files for tagging updated building locations to spatio-temporal movements of occupants. Moreover, for the modeling of occupant trajectories, the temporal resolution is kept maximum i.e. seconds, whereas the spatial resolution is kept to rooms.

Understanding occupant behaviors is a complex phenomenon. It not only involves the process of tracking movements with their dynamic context and computing the probabilities of their different actions inside the building that is one of the prerequisites of any kind of behavioral analysis but also needs to incorporate several external factors (e.g. weather information, data from building management and monitoring systems, etc.) to study occupant behaviors in more detail by

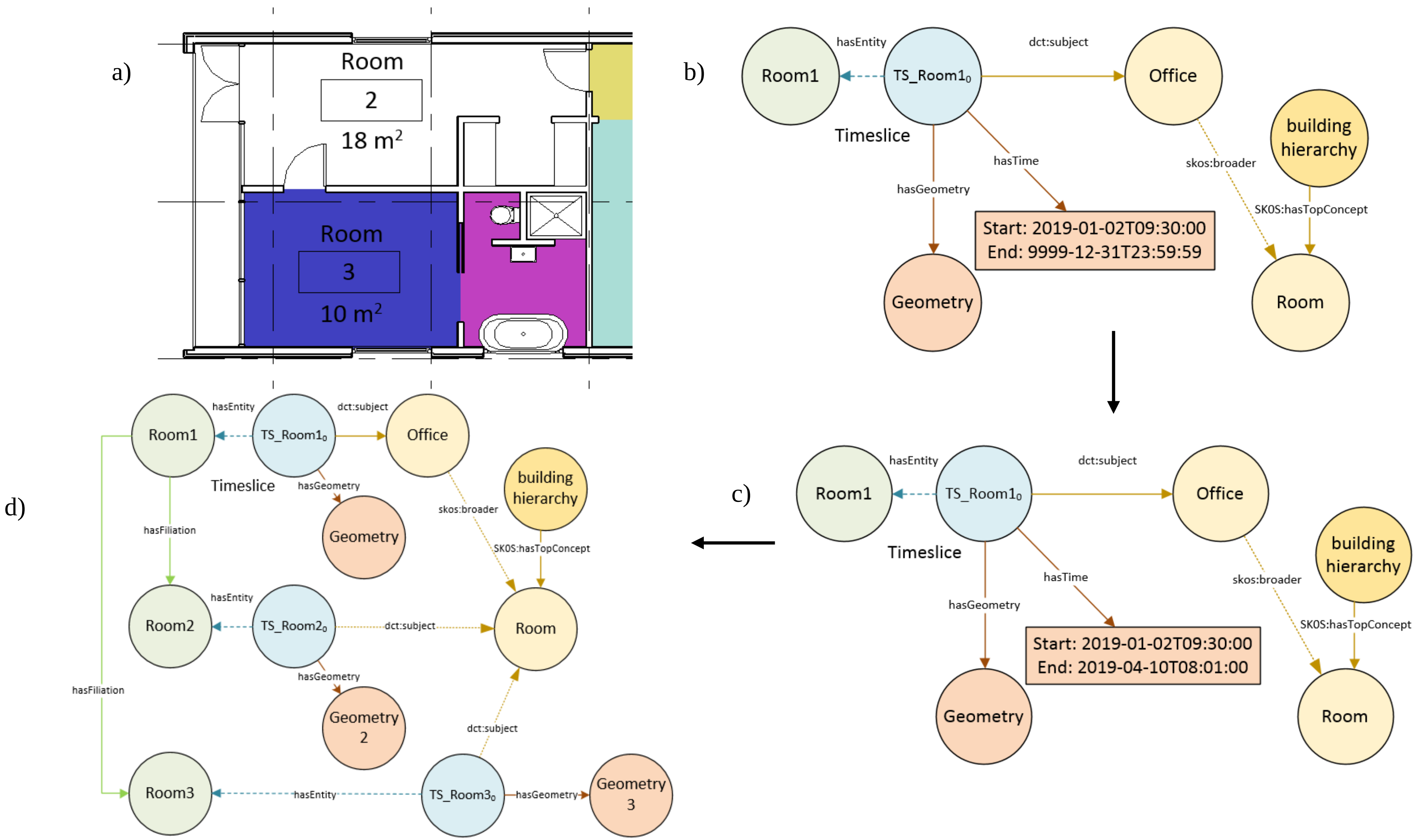

Fig. 3-10 Storing the information of newly created building rooms in a data model

extracting insights about occupant activities from their actions. However, a scope of the proposed framework is kept limited to include dynamicity of the building environment into occupant movements and presence by enriching a DNAS ontology (particularly actions) which can help to infer occupant activities for different facility management applications.

Chapter 4 - From sensor data acquisition to semantic trajectories

Outline

To study human behaviors, the capturing of the occupants` movements and presence were chosen. Later, BLE beacons were selected for analyzing the occupant movements using the OBiDE framework. First, preliminary measurement data analysis of beacon data was performed for achieving better accuracy in identifying the building locations. Second, the processes including a) trajectory data collection using BLE beacons, b) pre-processing as well as c) semantic enrichment of occupant movements with related contextual data of buildings for targeted case studies are described. In the end, a method is mentioned for detecting the errors in the BLE beacon trajectory data after capturing the occupant movements.

4.1 Sensor Data for Indoor Geo-localization

The effectiveness of a developed framework (see Fig. 3.4) in understanding occupant behaviors using spatio-temporal trajectories of occupants is completely dependent on the identification of the accurate building locations. Based on the benefits perceived from the literature on BLE technology (Paek et al., 2016), Bluetooth beacons were used for spatio-temporal data collection. The environment varies in buildings because of a different material of walls, the density of occupants, and the type of equipment present inside a building (Paek et al., 2016; Ke et al., 2018). The environmental factors of a building affect the Received Signal Strength Indicator (RSSI) values of beacons in identifying the precise locations of building occupants. To understand the building environment for deploying the beacons for achieving higher detectable RSSI values of beacons in the coverage areas, three different tests were performed which are described in Table 4.1.

Table 4-1 BLE behavior testing

Test	Purpose	Strategy
1	Understanding the degradation effect of the RSSI values in case of falling off the beacons from the wall.	Dropping a beacon from the height of the wall.
2	Simulating the concept of a dynamic building environment where the walls and infrastructure supports are added in building areas.	Mounting a beacon in a corridor and detecting its RSSI values inside the room to study the effect of the wall on a beacon.
3	Studying the effect of the existence of beacons tagged with two interconnected adjacent building locations for determining the correct location in a dynamic environment.	Mounting a beacon in a room and another beacon in a corridor and understanding the variation in the RSSI strength of both the beacons as the user moves across these locations.

For the 1st test, a beacon is mounted at 1.2 meter-height (see Fig. 4.1) and its RSSI values are recorded per second. The reason for selecting the height of 1.2 meters is to keep the height of a beacon identical at which the building users hold their handheld devices (Paek et al., 2016). In addition, the alignment of a beacon is adjusted to face it towards the position of a handheld device so that line-of-sight (LOS) communication is possible. By configuring the transmission power of a beacon, it is made detectable at a 5-meter distance. Using the settings mentioned above, an Android-based mobile application is used for detecting the beacon signals and recording its RSSI values per second while the distance between a beacon and a handheld device is kept constant i.e. 2 meters. However, this distance can be varied until the maximum distance where a beacon is detectable by a handheld device. Then, by keeping the same height, time duration and a beacon configuration, a beacon is dropped from the wall to the ground and then mounted it back to its original position for visualizing the amount of degradation in the RSSI values. The RSSI is measured in dBm (i.e. the power ratio in decibels between the measured power and a milliwatt) indicating the signal strength of a beacon and used for proximity analysis. The RSSI value of a beacon degrades as distance increases. The lower negative RSSI value in dBm shows the closeness of a beacon to a handheld device that is configured for detecting beacons (Paek et al., 2016).

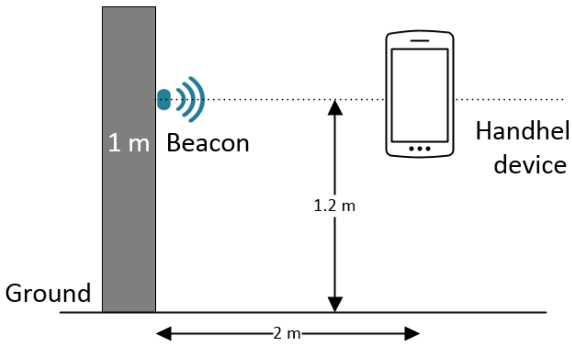

Fig. 4-1 Mounting a beacon at 1.2m-height

As shown in Fig. 4.2, the signal strength is decreased of approximately -25dBm when a beacon is dropped on the ground. The reason for decreased RSSI values is not only the falling off the beacon but also the interference caused by the user`s body who made a beacon to fall as it absorbs radiofrequency radiations. There will an obvious frequent case of falling off the beacons from their original mounting positions from the walls in dynamic building environments. The circumstances of falling off the beacons should have minimal effect on the RSSI values of the beacons for maintaining the required coverage area. Existing literature shows that signal smoothing techniques involving different filters (mean, median and Kalman) can be applied for reducing the effect of huge spikes and dips in the reported RSSI values (Zheng, 2015; Yan et al., 2013). The signal smoothing techniques are based on averaging algorithms that prevent the situations in the collected data where a beacon has reported a signal of -85 dBm and then it jumps the second to -60 dBm. For our study, a mean filter is used on the data to smooth the signals as it is efficient and robust in eliminating the outliers of lower deviations (see Fig. 4.2). However, advanced filters such as Kalman filters can also be applied as per the requirements and the degree of deviations in measured signals.

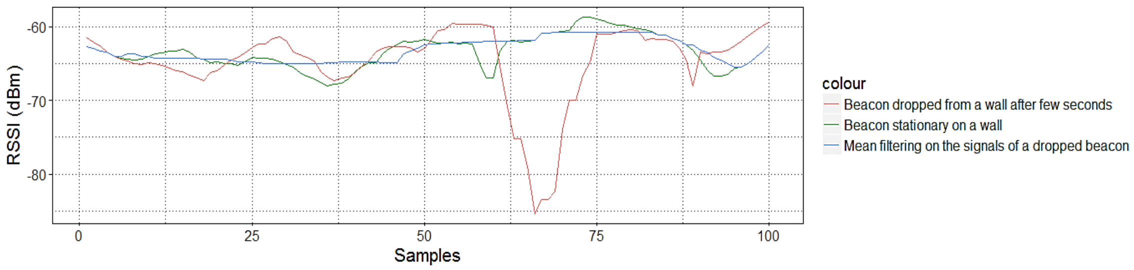

Fig. 4-2 Effect of falling off a beacon

For the 2nd test, a beacon is mounted in a room and its RSSI values are collected from inside the room (see Fig. 4.3). Later, to study the effect of a dynamic environment where walls are often added in the building infrastructure. For this, a wall is included in the area (i.e. Area24 consisting of a Room101 and a Corridor) covered by a beacon that is defined in the system for recognizing a building location. A beacon is placed outside the room in the corridor (see Fig. 4.4). Whereas, a room and a corridor belonging to the same location (i.e. Area24) defined in a data model. The RSSI values of a beacon mounted in a corridor are collected (see Fig. 4.5) from inside the room to study the effect of the wall on a beacon.

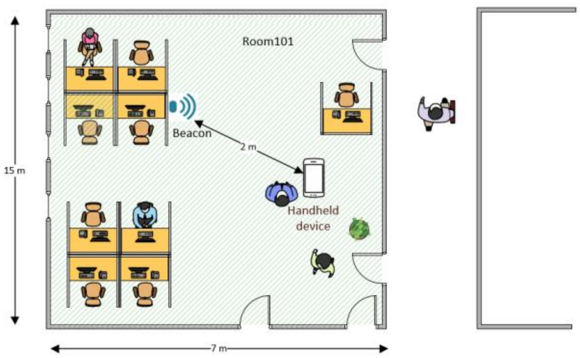

Fig. 4-3 No effect of an obstacle

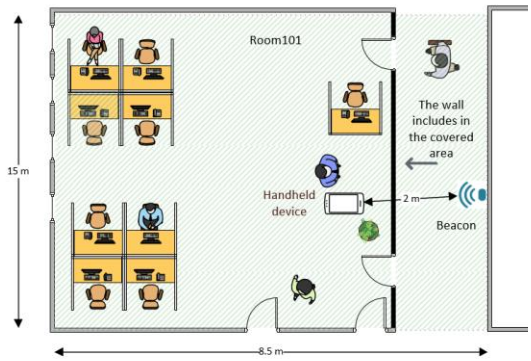

Fig. 4-4 Effect of a wall

As shown in Table 4.2, there is a reduction in the signal strength of around -12 dBm because of the interference caused by the wall.

Table 4-2 Effect of a wall on RSSI values

No.	Cases	Average RSSI value (dBm)
1	Beacon without any obstacle at a distance of 2m and height of 1.2m	-63.77
2	Beacon affected by a wall at a distance of 2m and height of 1.2m	-75.54

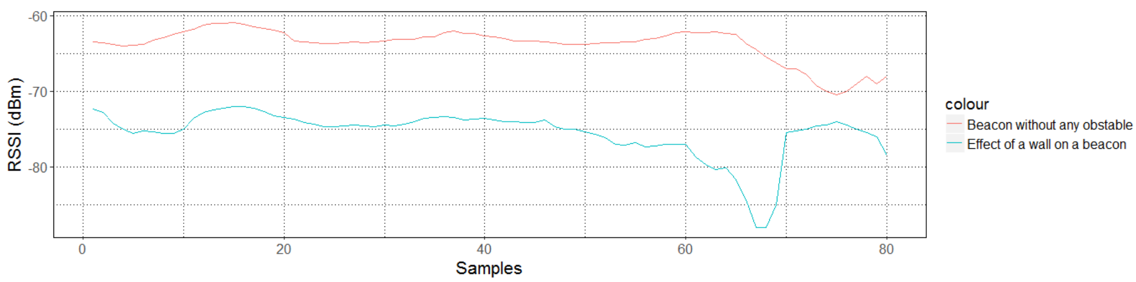

Fig. 4-5 RSSI plot to see the effect of a building obstacle (i.e. wall) on beacon

For the last test (as depicted in Fig. 4.6) to visualize the effect of the existence of beacons tagged with two interconnected adjacent building locations for determining the correct location in a dynamic environment. A beacon is mounted in a room named Room101 whereas, another beacon is mounted in a corridor. The transmission power of these beacons is configured in a way that two beacons are detectable in their neighboring locations. For RSSI data collection, a user with a handheld device moves away from Room101 towards a corridor approximately 1-m distance every second.

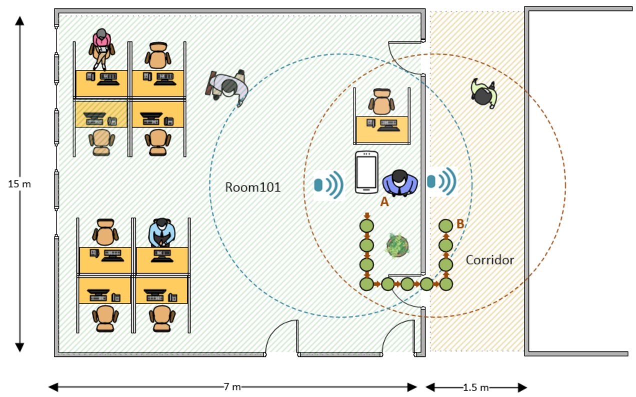

Fig. 4-6 RSSI measurement path in a dynamic environment

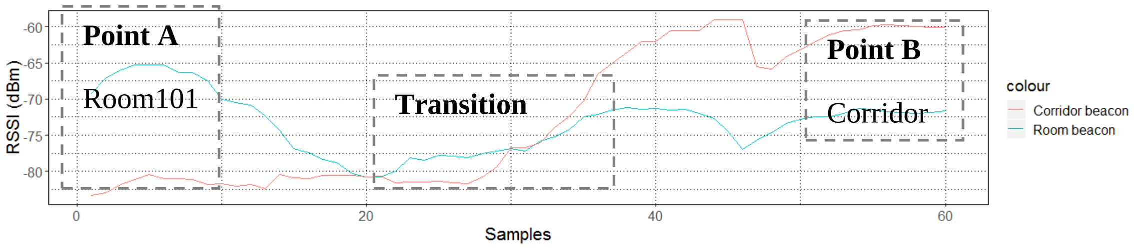

Fig. 4-7 RSSI measurement while a user moves from the point A to the point B

As shown in Fig. 4.7, when a user moves from room101 to a corridor, the RSSI values of both beacons' variates. Initially while at room101, the room`s beacon shows the stronger RSSI strength. As a user comes near the door and walks across the door connecting these two locations, the RSSI value of a room`s beacon starts getting degraded. As the user approaches near the corridor`s beacon, the signal strength of both the beacons gets inverted. Here, a point that is critical to understand is the effect of obstacles on the RSSI of beacons. The user stops its movements in a corridor from the same distance where a user has started its movement in a room i.e. 2m. The RSSI values of both beacons at a 2m distance should be nearly identically but that is not the case in Fig. 4.7. The reason for a lower RSSI value in a room is because of the deployment of other electronic equipment operating that must be interfering with the BLE signals as well as the presence of people who absorbs the radiofrequency radiations (Paek et al., 2016). While in a corridor, the closed space introduces the scattering effect in the beacon`s signal which ultimately increases the RSSI value of a corridor`s beacon.

The conclusion that is perceived from this test is that the existence of multiple beacons across different building locations does not affect in collecting the strongest RSSI values for detecting the most precise building locations of a user. However, collisions may occur if the material of walls is not dense enough which introduces the effect of attenuation in the RSSI signals. For our case study as shown in Fig. 4.7, the change of locations can be observed clearly while analyzing the RSSI values of beacons. This will eventually help us to collect precise proximity data of building users.

4.2 Indoor Localization Error Reduction Process

The approach that is used for reducing the proximity errors (resulted from the scenarios of tests 2 and 3 as discussed above) due to the obstacles in our study is to first capture data values in shorter time intervals and then applying the K-means clustering algorithm (Žalik, 2008) to identify the precise building location by computing the central points of the users in collected data values (see Fig. 4.8). The purpose of using this approach is to take a mean of location coordinates of detected locations than just relying on a signal beacon reading for identifying a building location. The central points of the user movements are later mapped with geographical coordinates of the building locations extracted from the OSM file to calculate the accuracy in recognizing the building areas where the users are present. The application of the K-means clustering algorithm along with a smoothing technique has enabled us to achieve a high location tagging accuracy.

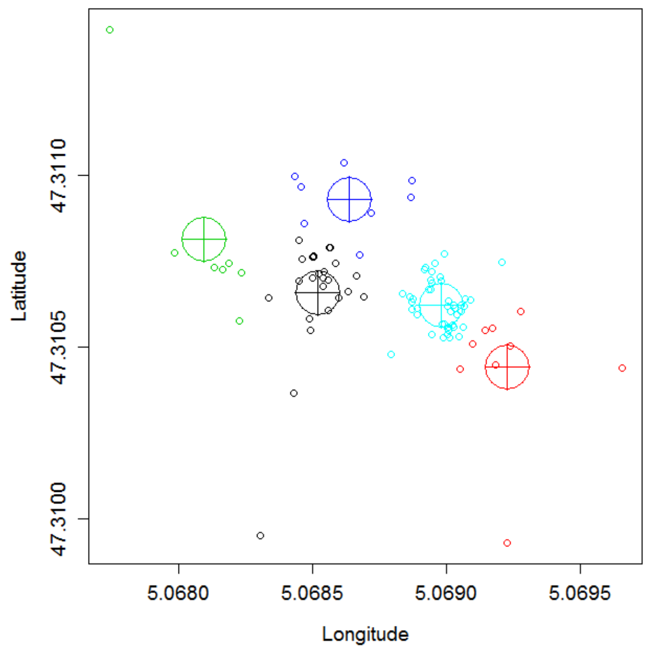

Fig. 4-8 Calculating centroids of user movements for precise location identification

4.3 Data Collection and Pre-processing

For collecting the spatio-temporal data, around 200 BLE beacons (see Fig. 4.9) were installed in two buildings (see Fig. 4.10). The information about buildings is described in Table 4.3. The deployment plan of the beacons is constructed using software to cover each building location with a range of at least three beacons by limiting each beacon`s signal strength to 5 meters. The

Fig. 4-9 Bluetooth beacons

software takes an OSM file of a building as an input, estimates the number of beacons required for maximum coverage and outputs the deployment plan (see Fig. 4.10) of the beacons according to the floor area of a building. The process is repeated for each floor of two buildings. Using the deployment plans, the beacons were mounted on different locations in buildings. If the BLE beacons are relocated due to the alteration of spaces and temporary structures in the dynamic environment, the positions of the beacons in the database file for tagging the locations against each longitude and latitude pair values should be updated accordingly. For acquiring the location coordinates (longitude and latitude pair values) of building users, an Android application (see Fig. 4.11) is installed in the handheld devices of building users. The application connects with a database containing information about the deployed beacons. In a database, each beacon is identified using its unique identification which corresponds to its assigned building location. As an application launches in a device, it detects the neighboring beacons and selects the best three beacons` signals to perform the geo-localization. Based on the received signals` strength of the beacons, location coordinates are generated by the application (see Table 4.4) and stored in a document database i.e. Mongodb for further processing. Later, these location coordinates are pre-processed in R studio. As real-life spatio-temporal data captured from different location acquisition devices very often suffer from noise and interferences because of the environment (Zheng, 2015; Yan et al., 2013). Loss of signals, sensor battery outage, and sampling misadjustments are some of the main reasons for having noisy spatio-temporal data (Zheng, 2015). For reducing the level of noise in the collected data, a median filter is used as it is robust and preferred for datasets having outliers with low deviations (Zheng, 2015)

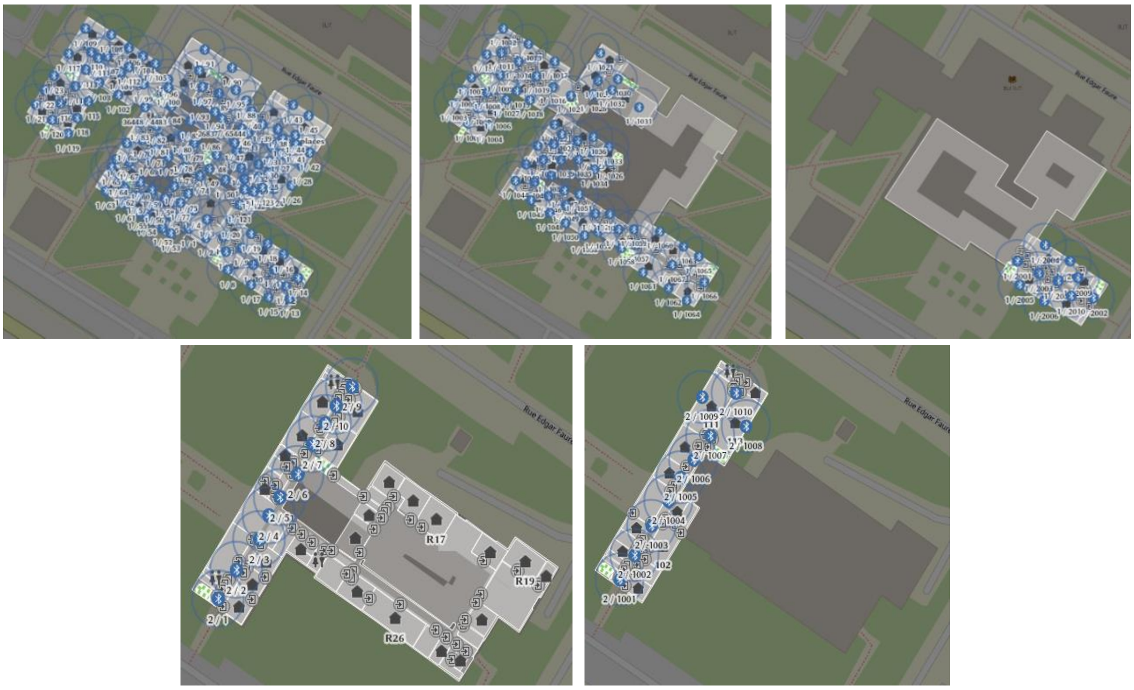

Fig. 4-10 Deployment maps of BLE beacons of different floors in two buildings

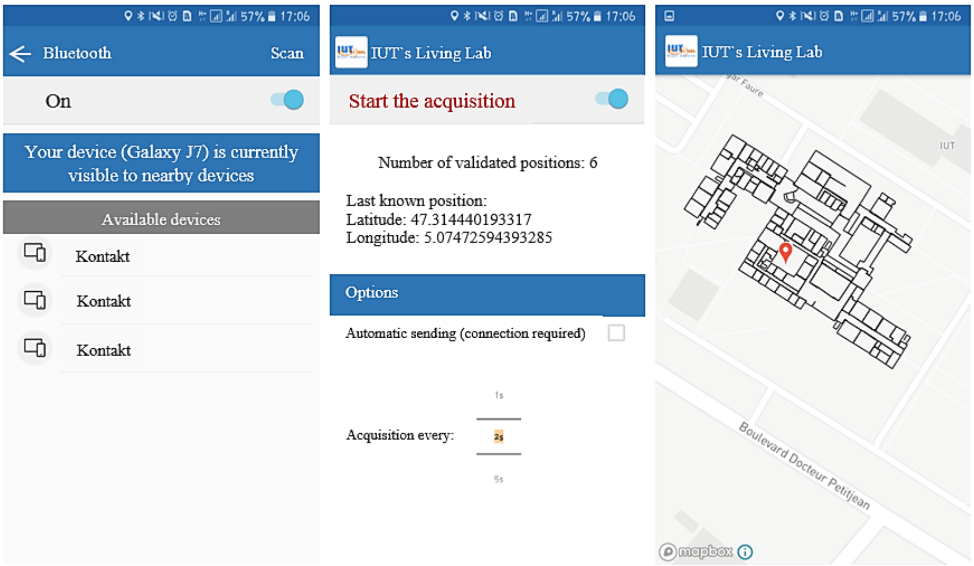

Fig. 4-11 Mobile application to collect beacon data for indoor geo-localization

.

Table 4-3 Building information

No. of buildings	Type of buildings	Occupant type	No. of users
2	Educational	Students	11

Table 4-4 Trajectory dataset information

Dataset	No. of trajectories	No. of location coordinates	Duration of study	Sampling frequency
Building users	30	8,426	2 weeks	5 seconds

4.3.1 Data cleaning

Data cleaning is the first step that R will perform on data retrieved from MongoDB. For this research, a median filter is used because of its robustness characteristic, whereas a mean filter is not recommended because of its high sensitivity to outliers (Zheng, 2015). 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 a sliding window mechanism that covers n temporally adjacent values of z_i as shown in the below-mentioned equation.

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

The choice of a median filter for data cleaning is also made because of the high-sampling rate of the location data; which makes the median filter a good option. Though, if a sampling rate of trajectories is too low than a median filter is not recommended and advanced filters such as Kalman filters can be considered for noise reduction. Figure 4.12 shows the actual and filtered trajectories of a worker.

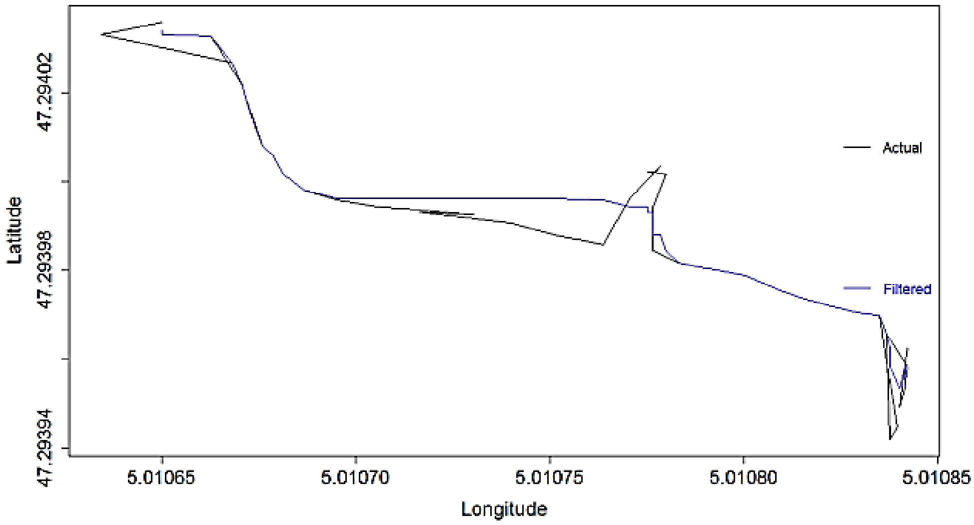

Fig. 4-12 Actual and filtered trajectory of a worker

4.3.2 Stay points detection

After cleaning the location trajectories, the data processing algorithms in the R environment will calculate the stay points of occupants using Algorithm 1. Stay points are the geographic locations where an individual has spent significant time within a certain distance. A stay location acts as a virtual location encompassing a set of successive location points (Zheng, 2015). The purpose of extracting the stay regions in trajectories is to identify the building regions where users are spending most of their time. This information will help in tracking the occurrence of unanticipated situations in the building if the stay duration is greater or less than the required. The calculation of stay points depends on two parameters, a distance threshold (D_{thresh}) and a time threshold (T_{thresh}). By setting the distance (D_{thresh}) value to 3 meters and time threshold (T_{thresh}) value to 20 minutes, stay locations in a trajectory are identified as shown in Fig. 4.13. The reason for setting up a 20-minute threshold is that time intervals smaller than 20 minutes possibly represent a user was only passing through the stay region for a short time period. Though, these thresholds (hyper-parameters) for identifying the stay regions are totally application dependent and can be changed according to an indoor or outdoor environment. Zheng (2015) proposed an algorithm for finding stay points in trajectories (Zheng, 2005). 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,

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

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

Algorithm 1. Stay Points Detection

```
Input: GPS Trajectory(G) ← {G_lat, G_lon, G_t}, Distance Threshold
(D_thresh), Time Threshold (T_thresh)
Output: Stay points (S) ← {S_lat, S_arrv, S_leav }
   1. i ← 0, Total number of GPS points (G_num) = |G|
   2. While (i < G_num)
   3. {
                 j ← i + 1
                 While (j < G_num)
                 {
                  Distance ← distHaversine {(G_lon[j], G_lat[j]), (G_lon[i], G_lat[i])}
                            If (Distance > D_thresh)
                            {
                            deltaT ← G_t [i + 1] − G_t [i]
                            If (deltaT ≥ T_thresh)
                                   {
                                   M_lon ← mean (G_lon[i], G_lon[i + 1])
                                   M_lat ← mean (G_lat[i], G_lat[i + 1])
                                   S_arrv ← G_t[i]
                                   S_leav ← G_t[i + 1]
                                   }
                            i = j; break;
   4.                }
   5.            j ← j + 1
   6.           }
   7. }
```

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). Algorithm 1 first checks if the distance between a point under consideration and its successors in a trajectory larger than a specified threshold. Then it 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, we have set D_{thresh} to 5 meters and T_{thresh} to 20 meters for stay points detection. There are four stay points detected in a worker A trajectory as shown in Fig. 4.13. However, the number of stay points can be increased or decreased as these are totally dependent on the values of a distance threshold (D_{thresh}) and a time threshold (T_{thresh}).

4.4 Tagging Locations

For tagging building locations to spatio-temporal trajectories, a process of semantic enrichment is used. Other than location-related characteristics of spatio-temporal trajectories, there exists an important characteristic of movements that is the variety of various travel means by which mobility has taken place. For example, for an outdoor scenario, 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

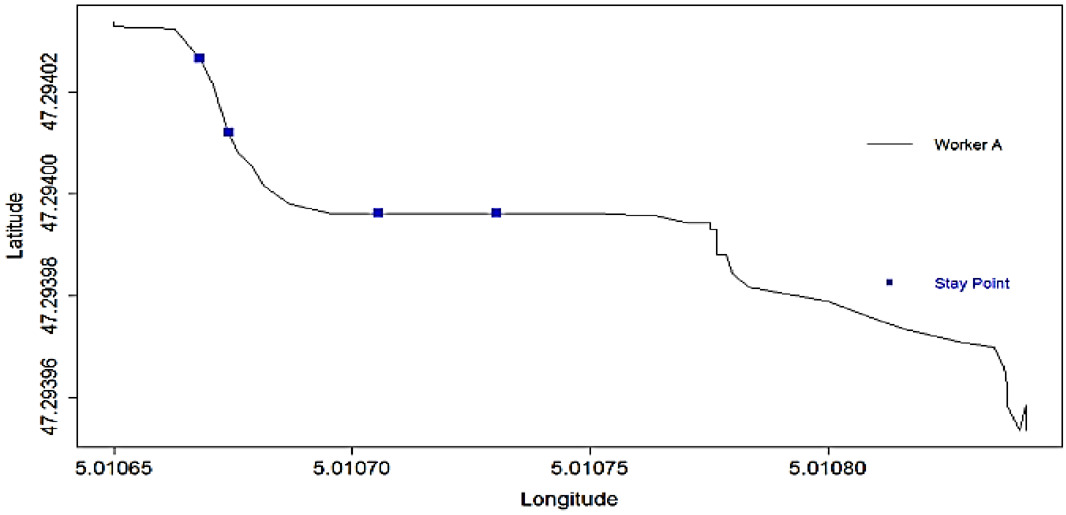

Fig. 4-13 Stay points detection in a worker`s trajectory

visualize people 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 spatio-temporal trajectories with such additional data is known as a semantic enrichment process (Yan et al., 2013). These additional data categorizing stops and moves, and distinguishing different types of moves are known as annotations that are attached to a trajectory either to some of its parts or as a whole (Yan et al., 2013; Zheng, 2005). 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 a semantically enriched trajectory could be the following (Yan et al., 2013).

(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, -)

The above example includes generic movement characteristics (e.g., stops and moves), application-specific geographical objects (e.g., office and work) and additional behavioral context (e.g., work) (Mousavi, 2016). Once trajectories are annotated and stored in a trajectory database, it can be used to detect and analyze trends and behaviors of moving objects. The 'STriDE' model is used for performing the semantic enrichment of collected trajectories. For tagging the semantic information (see Table 4.5) related to the building locations with the TSs, the STriDE model used ontologies which offer the conceptualization of a domain to show relations among different concepts. Here, 'concepts' define the different building locations as objects having a set of attributes (see Fig. 4.14). The STriDE model uses an OpenStreetMap building (OSM) file, a taxonomy and a set of rules written in the Resource Description Framework (RDF) language for performing semantic enrichment process. An OSM building file is in the Extensible Markup Language (XML) format that describes vector data for defining room boundaries along with their links with each other. The purpose behind an OSM file is to feed the vector data for describing an entire building structure in the trajectory data model (see Table 4.5). For labeling and clustering the key-value pairs present in an OSM file, a taxonomy is written as a hierarchy of ‘concepts’ (locations) written as RDF triples using Simple Knowledge Organization System (SKOS) vocabulary. Moreover, a set of rules are created in the form of a JavaScript Object Notation (JSON) file for establishing relations between OSM key-value pairs with the taxonomy. As a result, Java objects (see Fig. 4.15) having semantic information are generated and stored in a triplestore (i.e. Stardog) to achieve a complete representation of a building. One important feature of the STriDE model which is responsible for keeping the building evolution for semantic information is that, during the data modeling, the geometry of an object is defined outside the ‘main entity’ (see Fig. 4.15). Here, an ‘entity’ describes the identification of a building location. To store the historical changes occurring during a building evolution, the STriDE data model keeps tracks the spatial, spatial-temporal and filiation relations of building entities (locations and users). The spatial relation holds the information about how a user in a building is related to a reference building location. The spatio-temporal relation tracks how two building locations or a location and a user are linked with each other at the same time.

Table 4-5 Spatial information to tag trajectory point with a building location

Location	Room101
Geometry type	Polygon
Location coordinates [longitude, latitude]	([5.068456,47.31081], [5.068349,47.31084], [5.068287,47.31084], [5.068243,47.31082], [5.068198,47.31078], [5.068470, 47.3108])

SPARQL Results (returned in 47 ms)

userName	profileName	concepts
User 1	User profile	Building, Floor, Door, Corridor, Footway, Stairs, Room, Office, Meeting room, Amenity, Bathroom
Maintenance 1	Maintenance profile	Building, Floor, Door, Corridor, Footway, Stairs, Room, Office, Meeting room, Amenity, Bathroom, Storage

Fig. 4-14 Users and their profiles

```
stride:W235 a stride:Entity ;
  skos:prefLabel "Corridor of floor 1" ;
  rdfs:comment "" .

stride:GEO-W235-0 a geo:Geometry ;
  geo:asWKT "LINESTRING (20 85, 15 65, 15 50, 20 50, 16 50, 15 15, 20 10, 15 15, 30 15,
30 20, 60 15, 50 15, 50 20)"^^geo:wktLiteral .

stride:W235-0 a stride:TimeSlice ;
  stride:hasStartDate "2019-02-01T00:00:00"^^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. 4-15 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)

After generating the semantic trajectories (see Fig. 4.16) using the stored information (see Fig. 4.15 and Table 4.5) in the STriDE model, users` trajectories are represented as a series of timeslices (see Fig. 4.17) for two different applications, as discussed in the next chapters.

traj	userName	location
stride:TrajOfWorker2-1	Worker 2	Outdoor pathway
stride:TrajOfWorker2-2	Worker 2	Storage room
stride:TrajOfWorker1-1	Worker 1	Outdoor pathway
stride:TrajOfWorker1-2	Worker 1	Corridor of floor 0
stride:TrajOfWorker1-3	Worker 1	Office 1
stride:TrajOfWorker1-4	Worker 1	Corridor of floor 0
stride:TrajOfWorker1-5	Worker 1	Outdoor pathway
stride:TrajOfWorker1-6	Worker 1	Storage room

Fig. 4-16 Trajectories database having timeslices of two users

s	p	o
stride:TrajOfWorker1-3	rdf:type	stride:TimeSlice
stride:TrajOfWorker1-3	stride:hasStartDate	2018-01-02T09:06:00
stride:TrajOfWorker1-3	stride:hasEntity	stride:TrajOfWorker1
stride:TrajOfWorker1-3	stride:hasEndDate	2018-01-02T09:30:00
stride:TrajOfWorker1-3	stride:isTrajectoryOf	stride:Worker1
stride:TrajOfWorker1-3	stride:hasLocation	stride:W5

Fig. 4-17 Timeslice description of a Worker1 trajectory

4.5 Indoor Geo-localization Quality

4.5.1 Accuracy of location identification

The dataset for tests was composed of approximately 8,426 location points of 11 building users are collected using BLE beacons during different intervals in 2 weeks with a sampling frequency of 5 seconds. This data is generated by the University students in Computer sciences department. Then, the semantic enrichment system is used to transform the spatio-temporal trajectories into semantic trajectories by incorporating the contextual information about the building. The accuracy in tagging the building locations identified from the processed trajectories with the information residing in our data model was around 90% (see Table 4.6). Some of the building locations were not labeled correctly because of indoor interferences caused by the building objects that generated inaccurate recognition of the building locations using BLE beacons.

Here, the tagged building locations refer to the accuracy of spatially joining (i.e. $Trajectory\,point \bowtie_{\theta} Region$) the geographical coordinates of a user with the building information stored in the form of majorly as polygons using a set of geographical coordinates (longitude and latitude pair values) extracted from an OSM file of a building (see Table 4.5). Here, a parameter 'θ' is based on the topological spatial relation (e.g. distance). After computing the θ, spatial joins are performed with the centroid of user locations to extract the precise building location. More details on the spatial joins can be found in (Z. Yan, 2011) research.

Table 4-6 System accuracy for identifying locations

Detected user locations	Correctly tagged building locations	System Accuracy
150	136	90 %

4.5.2 Analyzing the positioning errors

To find out the errors (see Table 4.6) in the dataset collected from the BLE beacons which eventually caused the detection of locations of users in a building. The queries were executed for identifying certain movements of the users which are not correctly acquired by the system (see Fig. 4.18). For example, at a time $t = 1$, a user is detected in Room1. Whereas, at $t = 2$ a user is in Room2. However, these two rooms are not connected as can be observed using a BIM model. By bearing in mind the sampling interval (i.e. 5 seconds) of trajectory data, it is not possible to have such quick movement of a user between two locations that are not connected. For instance, a user moved from Room1 to Room2. Just for the sake of keeping an explanation simple for an example, 'days' are used as start and end dates of the timeslices. However, in our system start and end dates contain the precise date and time values. Using the trajectory data of a user, we know that a user moved on Wednesday (as stated by the end day of the 1st timeslice (i.e. TS_{T1}) and the start day of the second timeslice (i.e. TS_{T2}). For verifying the validity of this user movement, we need to find out the states of the rooms (Room1 and Room2) at the time of movement (i.e. somewhere on Wednesday). Here, the term 'state' refers to a timeslice holding an identification, alphanumeric properties, spatial representation, and temporal fields. At the occurrence of the changes in the alphanumeric and spatial (geometrical) properties of an entity (i.e. room), a new timeslice is created. For verifying the two connected timeslices, the end day of a timeslice should be the start day of the succeeding timeslice of a room. For tracking, we will check each generated timeslice of the rooms for finding the one for which the Wednesday is between the start and the end day. For Room2, it's easy to check as we have only one timeslice starting on Monday (see Fig. 4.18). To verify the connectivity of Room2 with a Room1, we will analyze the series of generated timeslices of Room1. For Room1, we can see that the 1st timeslice (TS_{R1}) starts on Monday and ends on Friday. Whereas, the 2nd timeslice ($TS_{R1`}$) which is resulted by the change in the geometry of a room starts on Friday. The day of movement which was 'Wednesday' is between Monday and Friday has no concern with the 2nd timeslice (i.e. $TS_{R1`}$) as this timeslice starts from Friday. From here, we know that the movement is done from the 1st timeslice (i.e. TS_{R1}) on Wednesday to Room2, eventually generating a timeslice TS_{R2}. But a timeslice TS_{R1} of Room1 is not directly connected to a timeslice TS_{R2} of Room2, as shown in Fig. 4.19. It means that Room1 was not initially connected with Room2. Over time, the geometry of Room1 is changed (generating a new timeslice i.e. $TS_{R1`}$) which might be the result of the construction of a new door or a pathway between Room1 and Room2. This change in the geometry makes Room1 connected to Room2 in the system. However, this was not the case in the observed trajectory data since the movement made from TS_{R1} is not possible, leading to an error (see Fig. 4.20 to see some of the errors). The errors in the system can be majorly caused because of the below reasons;

- Stopping the mobile application at a certain location and restarting it from another location can transmit the trajectory point of the previous location. This is usually caused by a delay in the detection and the transmission of a collected trajectory point.
- Some beacons might have fallen from their actual locations. The beacons were later tagged to the wrong places and can potentially cause errors during the trajectory data collection.
- The geometries of some locations are not fully incorporated in the model because of the new construction works. This might be the reason for incorrect tagging for the locations to the trajectories by our system.

However, for our scenario that is built for an already constructed building, the incorrect location tagging (see Fig. 4.20) in the trajectories resulted because of the former two reasons.

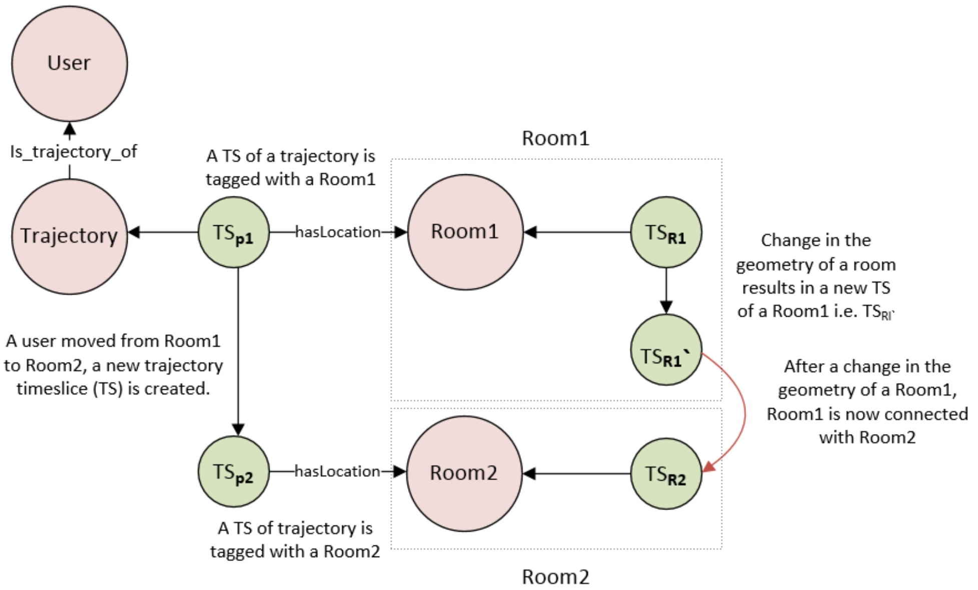

Fig. 4-18 A scenario for extracting the positioning errors

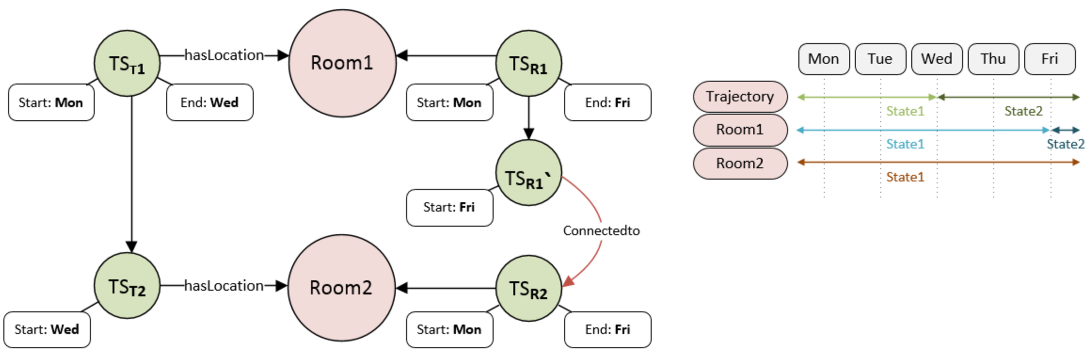

Fig. 4-19 An example of a scenario for extracting the positioning errors

SPARQL Results (returned in 1828 ms)

user	from	to
stride:b1bb9e247f060f5e	stride:W265	stride:W206
stride:b1bb9e247f060f5e	stride:W206	stride:W22
stride:b1bb9e247f060f5e	stride:W22	stride:W242
stride:b1bb9e247f060f5e	stride:W242	stride:W22
stride:b1bb9e247f060f5e	stride:W22	stride:W105

Fig. 4-20 Positioning errors showing user movements between not-connected locations

4.6 Discussion

The major contributions of this chapter were in two-fold;

(i) **Development of data collection and trajectory pre-processing subsystem:** As trajectory data holds multifaceted characteristics include: time (i.e., the position of the object on the timescale), position of the object in geographical coordinate system, direction of the object, speed of the object, change in the direction, acceleration (i.e., a change in speed) and distance traveled (Zheng, 2015). A system is developed to extract such characteristics to better understand the trajectories. After extracting the basic trajectories` characteristics, identification of the stay regions is achieved from workers` trajectories that will help in recognizing important regions in the building for categorizing movements.

(ii) **Development of semantic enrichment subsystem:** To enrich trajectories with contextual information, a data model named STriDE is used. The system has been designed to include application domain knowledge and geographic database information (OSM file) for transforming pre-processed trajectories into semantically-enriched trajectories.

After constructing the above-mentioned systems for trajectory pre-processing and semantic enrichment, the OBiDE framework is used for extracting the behavioral states (risky movements and unauthorized access) of the worker behaviors as discussed in Chapters 5 and 6. The framework is utilized for enhancing worker safety on construction sites. The reason for choosing the use-cases related to the construction site environment is to show the proof-of-concept working and feasibility of the proposed framework which is designed to hold evolving contextual information regarding buildings and the construction sites are the best example of such dynamic environments.

Chapter 5 - Worker movement analysis for safety management

Outline

Chapter 3 describes the OBiDE framework. This chapter is based on the 1st case-study to use the OBiDE framework for analyzing worker movements for construction safety management aiming to reduce the fatalities on construction sites. The contributions of this chapter are:

(i) ***Worker movements categorizing subsystem****:* After constructing the trajectory pre-processing (i.e. data cleaning and stay points detection) and semantic enrichment (using an OSM file of a building) system as described in Chapter 4, an entire set of trajectories belonging to a stay region (i.e. an ROI) is further analyzed by categorizing the worker movements into three states using the probabilistic HMMs along with the Viterbi algorithm.

(ii) ***BIM-based visualization subsystem:*** *T*he output of the Viterbi algorithm is visualized using a BIM model for identifying the most probable high-risk locations involving sharp worker movements and rotations.

5.1 Scenario and the Prototype System

Existing literature (Ilkovičová et al., 2016) shows that the movements of the occupants and the moving objects can be studied if step lengths and turning angles are periodically measured from their spatio-temporal trajectories. For monitoring the worker movements in a building, an OBiDE framework designed for dynamic environments is applied (see Fig. 5.1 and Fig. 5.2).

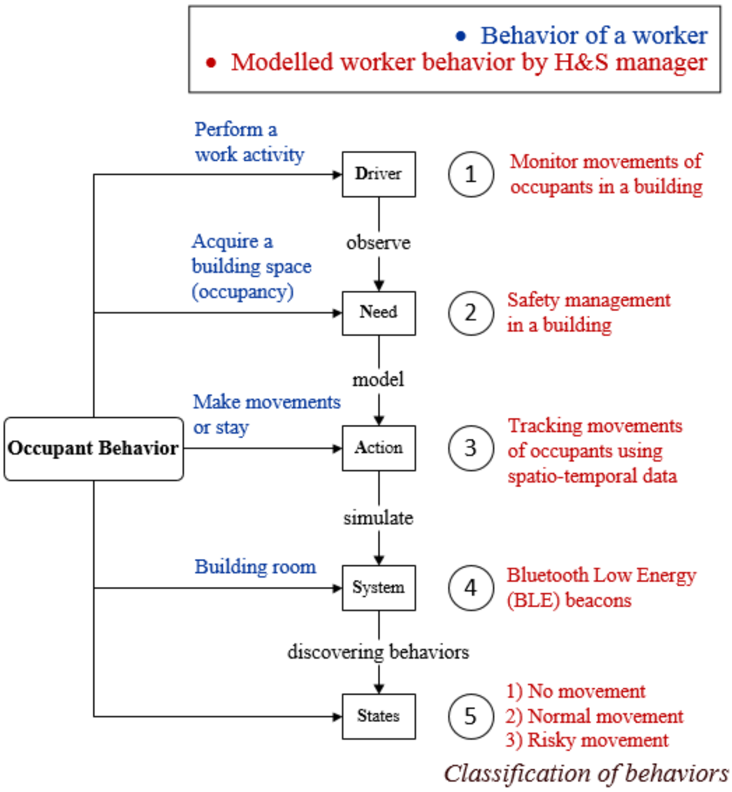

Fig. 5-1 Description of a scenario using OBiDE framework

In this case,

1) ‘driver’ is monitoring movements of occupants,
2) ‘need’ is achieving safety management in a building by identifying unsafe movements,
3) ‘action’ is tracking movements of workers using their spatio-temporal trajectories,
4) ‘system’ corresponds to BLE beacons for sensor data acquisition deployed in a building, and
5) ‘states’ are; a) no movements, b) normal movements having short steps and few turnings, and c) risky movements having long steps and many turnings.

Though, before categorizing workers` movements, specific regions in a building where the occupants stay for a longer duration need to be identified. These stay regions (Zheng, 2015) which are termed as Regions of Interest (ROIs) in our study (as discussed in Chapter 4) are more critical to monitor for extracting trajectory movement insights than moving locations as the occupants are spending most of their time there. To achieve this, a scenario is initially created (see Fig. 5.2). The purpose of using a scenario-based methodology is because the scenarios are

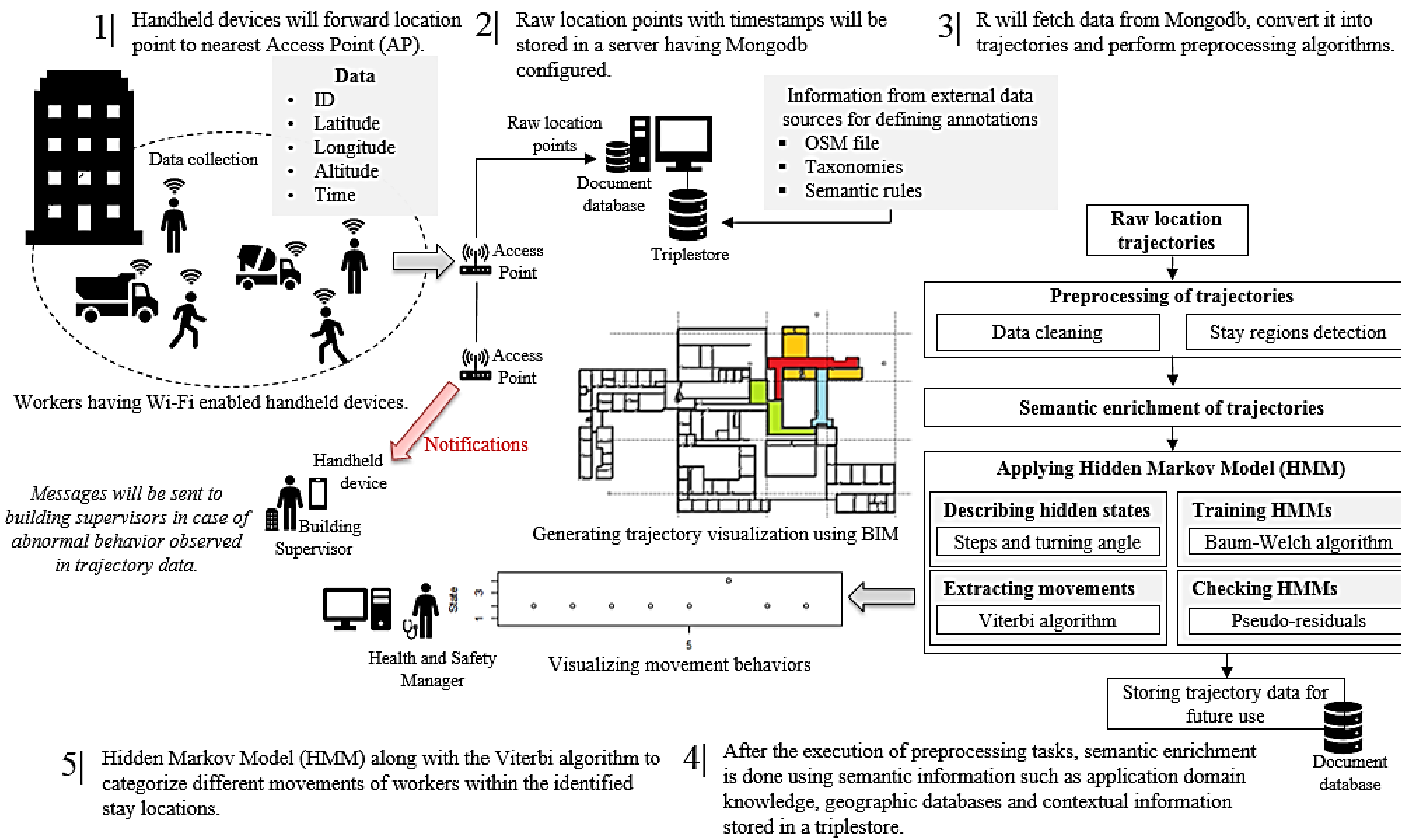

Fig. 5-2 Prototype system scenario

built on qualitative causal thinking which facilitates to communicate (Van der Heijden, 2011) and brings project participants (developers, building supervisors, H&S managers, etc.) together towards a shared understanding of the situation i.e. monitoring the buildings for identifying the unsafe behaviors of the construction resources (workers and machinery) (Carroll, 1995).In the created scenario (see Fig. 5.2), capabilities of BLE beacons, a semantic enrichment technique for incorporating the dynamicity of construction environments and a statistical method for categorizing the worker movements are mapped together to construct a real-time monitoring system for improved worker safety. Based on the scenario (see Fig. 5.2), a prototype system named ‘WoTAS’ (Worker Trajectory Analysis System) is developed by considering the roles of building supervisors and H&S managers. First, spatio-temporal trajectory data of building users working inside the already constructed building is acquired for experimental analysis. Then, pre-processing is performed on the raw trajectories and stay regions of the users are identified inside the building (as described in Chapter 4). Each stay region consists of a set of trajectory segments of the users. For enriching trajectories with the related contextual information using external data sources both openly available and private data related to the buildings and the construction sites, a process of semantic enrichment is executed. As discussed in Chapter 4, using the stored contextual information from a triple store, the semantically-enriched trajectories are generated (see Fig. 5.3 and 5.4). The semantically-enriched processed trajectory is later used for the extraction of movement-related behavior information by applying the HMM probabilistic framework.

SPARQL Results (returned in 55 ms)

traj	userName	location
stride:TrajOfMaintenance1-1	Maintenance 1	Outdoor pathway
stride:TrajOfMaintenance1-2	Maintenance 1	Storage room
stride:TrajOfUser1-1	User 1	Outdoor pathway
stride:TrajOfUser1-2	User 1	Corridor of floor 0
stride:TrajOfUser1-3	User 1	Office 1
stride:TrajOfUser1-4	User 1	Corridor of floor 0
stride:TrajOfUser1-5	User 1	Outdoor pathway
stride:TrajOfUser1-6	User 1	Storage room

Fig. 5-3 Semantic trajectories of two users

SPARQL Results (returned in 47 ms)

s	p	o
stride:TrajOfUser1-1	rdf:type	stride:TimeSlice
stride:TrajOfUser1-1	stride:hasStartDate	2018-01-02T09:00:00
stride:TrajOfUser1-1	stride:hasFeature	stride:TrajOfUser1
stride:TrajOfUser1-1	stride:hasEndDate	2018-01-02T09:05:00
stride:TrajOfUser1-1	stride:isTrajectoryOf	stride:User1
stride:TrajOfUser1-1	stride:hasLocation	stride:W1

Fig. 5-4 Timeslices of a trajectory; each consisting of an identity for its representation, valid period and a geometric component for the spatial representation

5.2 Trajectory Insights

The main research objective of this study is to gain trajectory insights by recognizing the building areas where the unsafe worker movements are occurring that can potentially lead to accidents. As for recognizing and categorizing the movements, many case studies are present in the literature based on machine learning algorithms using Bayesian dynamic models (Wang et al., 2017) and clustering techniques (Seo, et al., 2015; Liu et al., 2015), state-based models such as simple Markov chains and HMMs (Rabiner, 1989), patterns matching algorithms and deep learning-based techniques (Fang et al., 2018). Though, HMMs-based method is chosen for our study because statistical HMMs are applied widely in many works (e.g. Fang et al., 2018; Michelot et al., 2016) for categorizing the trajectory movements and extracting patterns. However, advanced statistical techniques based on deep learning using neural networks (Fang et al., 2018) can also be applied for our scenario. The justification of choosing the HMMs is that, in the literature the human mobility is described as a series of Markovian stochastic processes where the probability distribution of a future state (i.e. safe or unsafe behavior or a next location) of a stochastic (i.e. a random) process (a trajectory in our case) is only dependent on its current state or a present location which eliminates the need of incorporating the preceding states and eventually minimal training data is required. The HMMs appear best for our application as we have a limited trajectory data captured in a duration of two weeks. Whereas, advanced deep learning-based methods are preferred when hundreds of thousands or even more records are available in a dataset for the training the model for generating the categorized movements (Fang et al., 2018).

5.2.1 Overview of Hidden Markov Models (HMMs)

Existing studies (Fang et al., 2018; Michelot et al., 2016; Hévízi et al., 2004; Li et al., 2016) consists of many solutions for describing the object behavior in time and the Markov Chain (MC) model is one of them. The MC is a state-space model in which the probability of having the next state depends only on the present state (Hévízi et al., 2004). The HMMs are the direct generalization of MCs which assumes that we cannot directly observe the exact states of the system and only a stochastic function of these states is visible to us (Rabiner, 1989). Generally, HMM has three main properties that define it (Rabiner, 1989). First, it assumes that the observation was generated by a process at time t whose state S_t is hidden from the observer. Secondly, 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 for predicting the future of the process, the hidden state encapsulates all the information we need to know from the process history (Rabiner, 1989). Thirdly, the supposition of the HMM is that the hidden state variable can only take T integer values $\{1, 2, 3, \ldots, T\}$. In general, an HMM λ is described by a set of three parameters (Rabiner, 1989) which can be written as the 3-tuple $\lambda = (A, B, \pi)$. Whereas A is the transition probability matrix, B is the emission probability matrix and π is the vector of the initial state probabilities.

HMMs are used extensively in real-world applications to study the behavioral patterns of moving objects (Wang et al., 2018; Du et al., 2018; Pastell et al., 2018; Dong et al., 2017; Postawka et al., 2017; Ronao et al., 2017; Williams et al., 2017; Ulmeanu et al., 2017; Zin et al., 2016). Numerous studies have been conducted based on the smart cameras, wearable sensors, and GPS devices to collect location data to track people, and other moving objects to investigate their different movement behaviors for better decision making (Fang et al., 2018; Wang et al., 2018; Du et al., 2018; Pastell et al., 2018). Understanding the building occupancy patterns, abnormal behavior identification and detection of dangerous activities of people, behavioral surveillance system for elderly care-and safety risk assessment on construction sites are the major HMM applications using the location data. For outdoor application scenarios, Global Navigation Satellite System (GNSS) technology and static cameras are widely used for locating people, where potential users were asked to enable GPS functionality of their handheld devices (Li et al., 2016). The use of GPS technology is low-cost as the necessary hardware is already present in most mobile devices. For developing indoor systems for HMM-based behavioral analysis, wireless networks such as wearable sensors for real-time and duration monitoring, accelerometer and gyroscope sensors for recognizing different human activities, CO_2, humidity, temperature, motion and smart electric metering sensors are used for inferring building occupancy (Chaney et al., 2016; Šabata et al., 2016; Liisberg et al., 2016). For an indoor building environment, a high cost will be associated with mounting sensors in the building environment. It is directly dependent on the size of the coverage area and the type of functionalities required from the sensing devices. However, the data acquisition from the sensor networks will be inexpensive as the wireless communication infrastructure already exists in most buildings. In the case of video data collection, the deployment of cameras will require extra hardware, as well as advanced image processing techniques will be needed to provide higher accuracies in the data acquisition and prepare the datasets for HMM-based analysis.

From the literature review provided above, it is noted that communication technology for capturing occupants` data should be carefully chosen before feeding it to the HMMs for categorizing the behaviors. A selection must be made as to what is more crucial for designing a sensor network for a particular behavioral application that should be low cost, high precision and having low to moderate complexity in deployment and management. In all the applications discussed above, there have been three fundamental scenarios for which HMMs are used (Rabiner, 1989). 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 $O = \{O_1, O_2, O_3, \dots, O_T\}$.

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.

5.2.2 Semantic trajectory insights using HMMs

For categorizing the occupant trajectory into different movement states for each ROI, an independent HMM is initially trained. According to the existing literature (Ciabattoni et al., 2019; Ilkovičová et al., 2016), the movement behavior of a moving object is defined by calculating individual step length and turning angle. However, before imputing the trajectory sequences into an HMM, we need to visualize the data to set the hidden state values using measurement variables. For our case, we have used values of 'step length' (i.e. the distance between two trajectory points) and 'turning angle' (i.e. the change in direction in radians from the previous point to the current point), which were extracted during the pre-processing of the trajectories. Step length (l_t) is calculated using the Haversine distance formula (Ilkovičová et al., 2016) between the locations (x_t, y_t) and (x_{t+1}, y_{t+1}) as below;

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

Where "r" is the radius of the earth. The reason for using the Haversine distance formula is because it is the most adequate method that calculates the geographic distance between two points on a sphere (Ciabattoni et al., 2019). 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 $[t, t+1]$ (Ciabattoni et al., 2019). For visualizing the occurrences (i.e. the frequency) of different values of the step length and turning angle in a trajectory during the sampling interval of 1 minute, time series and histograms are plotted in Fig. 5.5 to 5.8.

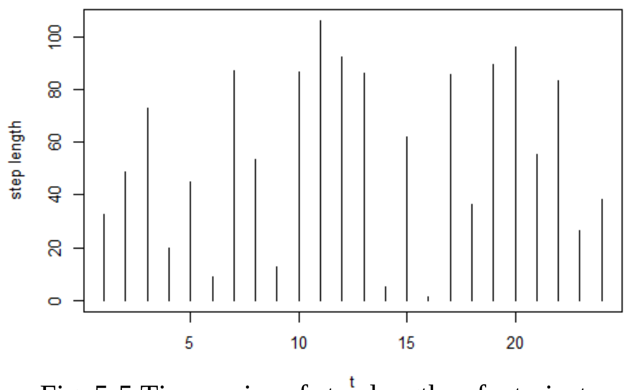

Fig. 5-5 Time series of step lengths of a trajectory

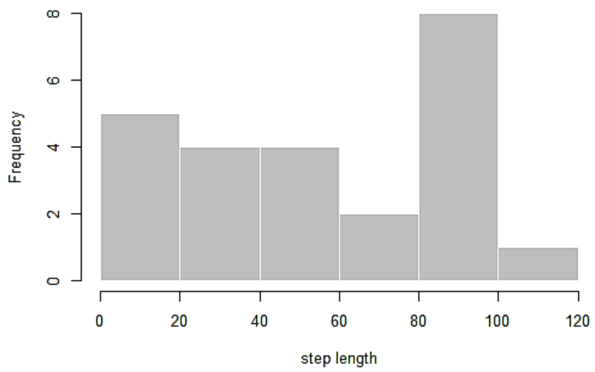

Fig. 5-6 Histogram of step lengths of a trajectory

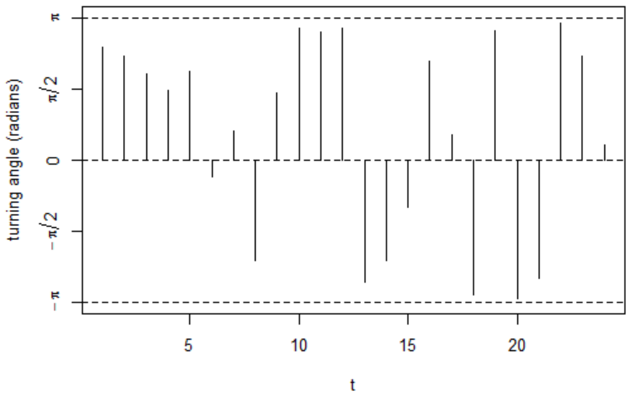

Fig. 5-7 Time series of turning angles of a trajectory

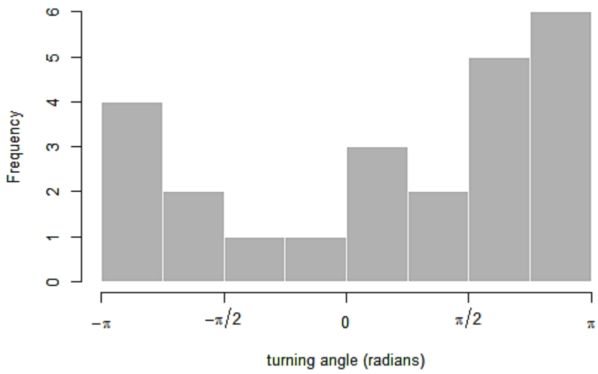

Fig. 5-8 Histogram of turning angles of a trajectory

The above-mentioned time-series plots (see Fig. 5.5 and 5.7) will help us to detect any outliers or sudden shifts in the trajectory. Whereas, histogram plots (see Fig. 5.6 and 5.8) provide us the frequency of the values of step length and turning angle in an entire trajectory. For movements' categorization of the workers within an ROI, we need to select the hidden states. This is the most important step while training the HMMs for generating the most probable values of the hidden states. The normal walking speed for an adult range from 1.0 to 1.6 meters per second (m/s) (Ilkovičová et al., 2016). However, by considering an indoor environment, 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. Using this method, three hidden states are defined using different values of 'step length' and 'turning angle' which are;

1. static (no movement),

2. normal movement ($0 < steps \leq 84$ and $\pi/2 \leq angle < \pi$), and

3) risky movement ($steps > 84$ $angle \geq \pi$)

However, the number of states can be decreased or increased as per the application requirements. The purpose of defining such hidden states is to identify unsafe worker movements within a stay region. For defining the hidden states in the HMM, we need to specify distributions according to the acquired 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 (see Table

5.1). For more details on these distributions, see Abramowitz and Stegun, (1972), and Gatto and Jammalamadaka, (2007) research. After defining the hidden states, the Baum-Welch algorithm (Rabiner, 1989) in the R studio is implemented which allows learning of the HMM parameters using the maximum likelihood approach and computes λ^* that maximizes the likelihood of the sample of training sequences $\chi = \{O^k\}_k^K$ namely $P(\chi|\lambda)$. Though, before training the HMMs, we must input initial probabilities. For our case, as already discussed above, there are three different states are denoted as $S_1, S_2, and\ S_3$. Dividing the total probability equally for all the states. This will give us initially estimated states` probabilities as mentioned below;

$$\pi = \pi_1, \pi_2, \pi_3 = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$

Table 5-1 Used distributions for defining states

Distribution	Parameters	Link function
Gamma	$mean > 0$	Logarithmic
	$standard\ deviation > 0$	Logarithmic
Von Mises	$mean\ \epsilon\ \{-\pi, \pi\}$	Logarithmic
	$concentration > 0$	Tangent $\left(\frac{mean}{2}\right)$

To fulfill the requirements of real construction processes, initial probabilities should be calculated based on the existing situation of construction sites and using historical records for improved accuracy in prediction methods. Out of 30 trajectories (as discussed in Chapter 4), 25 trajectories of randomly selected users were used for training the HMM. Whereas, 5 trajectories were used for validation purposes.

5.2.3 Model checking

After training the HMM using 25 trajectories of users, the trained model is evaluated against the training dataset. To evaluate the general goodness of the 3-state trained HMM, pseudo-residuals (Zucchini, 2000) (also known as quantile residuals) are calculated using a trajectory sequence $O = \{o_1, o_2, o_3, \dots, o_T\}$ for verifying whether the trained HMM is a true data generating process of a user trajectory or not. The residuals are what is left over after fitting a trained model and their values are equal to the difference between the trajectory observations and the corresponding fitted values (Zucchini, 2000). If the trained HMM model fits the data well, the data points in the qq-plot will be closer to the straight line and deviations of the points from the normality will indicate a lack of fit. The pseudo-residuals of the 3-state HMM fitted to the trajectory data of an occupant with the theoretical quantiles on the horizontal axis are displayed in Fig. 5.9. As shown in Fig. 5.9, the trained HMM is nicely fitted for the observation trajectory dataset and has few deviations from the straight line (i.e. the normality). However, the degree of goodness of a model to incorporate all the possible variations in the trajectory data can be improved by increasing or decreasing the number of hidden states (Bozdogan, 1987). For more details on pseudo-residuals, see Zucchini, (2000) research.

5.2.4 State decoding for model validation

Once the HMM is fitted using the training trajectory dataset, different test sequences from the trajectories of users are classified into three states which are S_1 , S_2 and S_3 using the Viterbi algorithm.

The Viterbi algorithm performs a global decoding process and generates the most probable sequence of states which have produced the observations under the trained HMM model (Rabiner, 1989). As an example, Fig. 5.10 shows the classification of one of the test sequence $O = \{o_1, o_2, o_3, \ldots, o_T\}$. The resulted states` sequences (one of them is shown in Fig. 5.10) are compared to the ground truth data which is collected by manually extracting the stays, normal and risky movements of the users from their trajectories using the values of step length and turning angle. Using these parameters, ground truth data is manually constructed. Later, a confusion matrix (Sokolova and Lapalme, 2009) is computed by inputting the predicated states of the trained HMM and the ground truth states` data for describing the performance of the trained HMM classification model. Finally, a precision-recall analysis (Sokolova and Lapalme, 2009) is conducted after computing the confusion matrix for determining the reliability of the trained model. The precision is defined as the ratio of how much of the predicted data is correct. Whereas, recall is the ratio of how many of the actual trajectory states were predicted (Sokolova and Lapalme, 2009).

Table 5-2 Confusion matrix for a 3-state HMM evaluation

No. of test samples = 2,132		Predicted States		
		$\boldsymbol{S_1}$	$\boldsymbol{S_2}$	$\boldsymbol{S_3}$
True States	$\boldsymbol{S_1}$	1066	23	0
	$\boldsymbol{S_2}$	30	810	75
	$\boldsymbol{S_3}$	6	17	105

$\mathbf{S_1} = \text{Stay}, \mathbf{S_2} = \text{Normal movement, and } \mathbf{S_3} = \text{ Risky movement}$

Using the R library of Caret package (Kuhn, 2019), the actual and predicted values were inputted, and the precision-recall parameters were extracted. The overall precision of the model was 83%, whereas the recall of the model was 89%.

5.2.5 Visualizing movement states using BIM

For the sake of the proof-of-concept demonstration of HMM-based analysis, a single trajectory is decoded into three different states (S_1 , S_2 and S_3) which corresponds to the movements of a user in a building in Fig. 5.10. The decoded trajectory needs to be linked with the building location from where it was captured. For visualizing the classified movement states resulted from an HMM by analyzing a user trajectory, a Building Information Modeling (BIM)-based software i.e. Autodesk Revit Architecture is used. A room in a Revit denotes a three-dimensional volume for representing a real building space (Autodesk Dynamo, 2019). Building spaces that need to be identified as individual rooms (POIs) should be properly bounded by the walls before placing ‘room tags’ on them. These tags are called ‘annotations’ and can be altered manually. Each

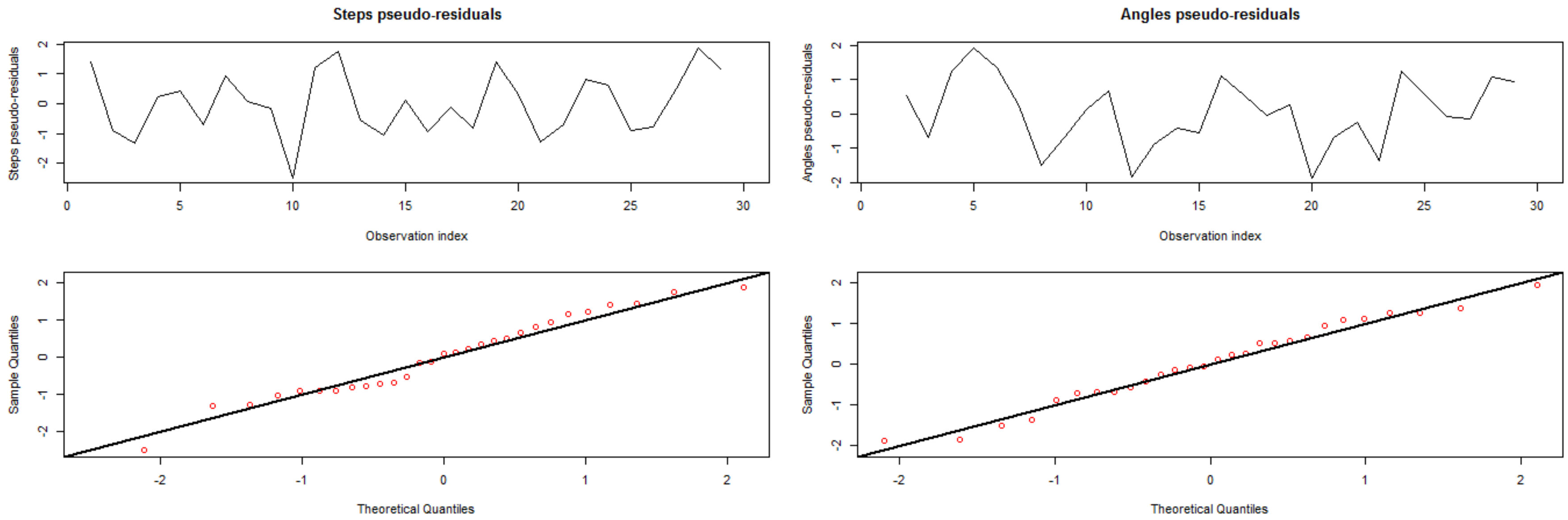

Fig. 5-9 Time series (top row) and qq-plots (bottom row) of the pseudo-residuals of the 3-state HMM

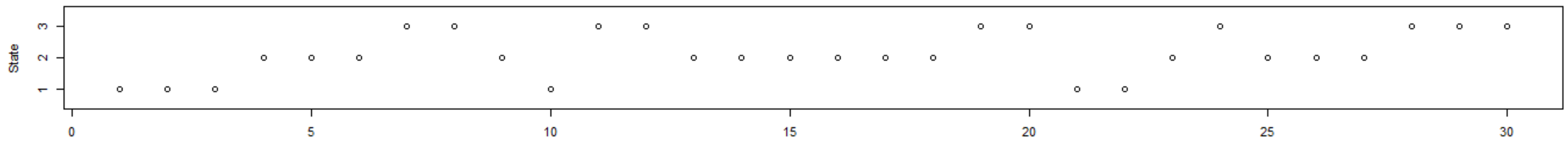

Fig. 5-10 Decoded sequence of states (S_1 , S_2 and S_3) of the trajectory observations of a user

tagged room in a Revit contains the semantic information in the form of a collection of parameters such as room number, room name, physical area, etc. used for viewing or editing the rooms. In our case, the parameter ‘room name’ acts as a unique identification for each building space representing a POI for visualizing the problematic locations in the Revit model using the output of the Viterbi algorithm. The naming convention of Revit rooms is set according to the tagging of the POIs as described in the STriDE model. For understanding the users` movements in the context of a building model, the input trajectory streams of the Viterbi algorithm are semantically enriched with their corresponding POIs` information which is stored in the STriDE model. For visualizing the decoded states of a trajectory on a BIM model, a Revit plug-in named ‘Dynamo’ (Autodesk Dynamo, 2019) is used. Dynamo is a visual programming tool that gives the ability to define pieces of logic using visual scripting by minimizing the requirement of extensive programming. Dynamo enables us to construct visual programs (known as graphs) by connecting ‘nodes’ with ‘wires’ (also called as connectors) for specifying the logical flow of information (Autodesk Dynamo, 2019). A Dynamo graph constructed for ‘WoTAS’ consists of three main steps (see Fig. 5.11).

As shown in Fig. 5.11, at first, a list of all the locations of a building model that were tagged as ‘rooms’ in a Revit software is extracted into Dynamo. This list is compared with the information of building locations having risky movements obtained from the excel sheet that is resulted by our prototype system. The risky movements are quantified based on their percentage of occurrences in a trajectory. For instance, in Fig. 5.10, a trajectory having 30 sample points is shown which spans over 2.5 minutes. An HMM has categorized this trajectory into 3 movement states. Out of 30, 10 states were identified as ‘risky states’. The percentage frequency of risky states is computed by dividing the number of occurrences of risky states by the total number of trajectory points and multiplying it by 100 i.e. $\left(\frac{10}{30} \times 100\right)$, a value of 33 is achieved for visualization. To quantify the calculated percentage, three different colors are used for generating the visualization which are; Red, Orange and, Yellow (see Fig. 5.12). The color scheme is designed using the OSHA color code standard for safety management at work (OSHA, 2019). The red color is used for visualizing such locations which have the highest percentage of risky movements i.e. 70 or more whereas, percentages less than and equal to 30 are shown in yellow color, and percentages greater than 30 and less than 70 are shown in orange color. The range of values for defining different colors was constructed for the sake of demonstrating the functionality of a prototype system to show the criticality of locations by changing their colors on a BIM model. These values can be altered accordingly based on user preferences. To show a process of decoding and visualization using a BIM model, a single user trajectory of a short time duration is used. However, in the case of multiple and long trajectories of several occupants that are collected throughout a day, the process will remain the same as described above but will be repeated for each individual trajectory. In the end, to assign the color to a location, a function of the averaging needs to be performed on the calculated percentages to acquire a unique value for representing different trajectories on a BIM model. In the 3rd step of our Dynamo graph, conditions for the color assignment are written for visualizing different types of risky locations on a BIM model. The changes in the color of the Revit rooms are achieved using the

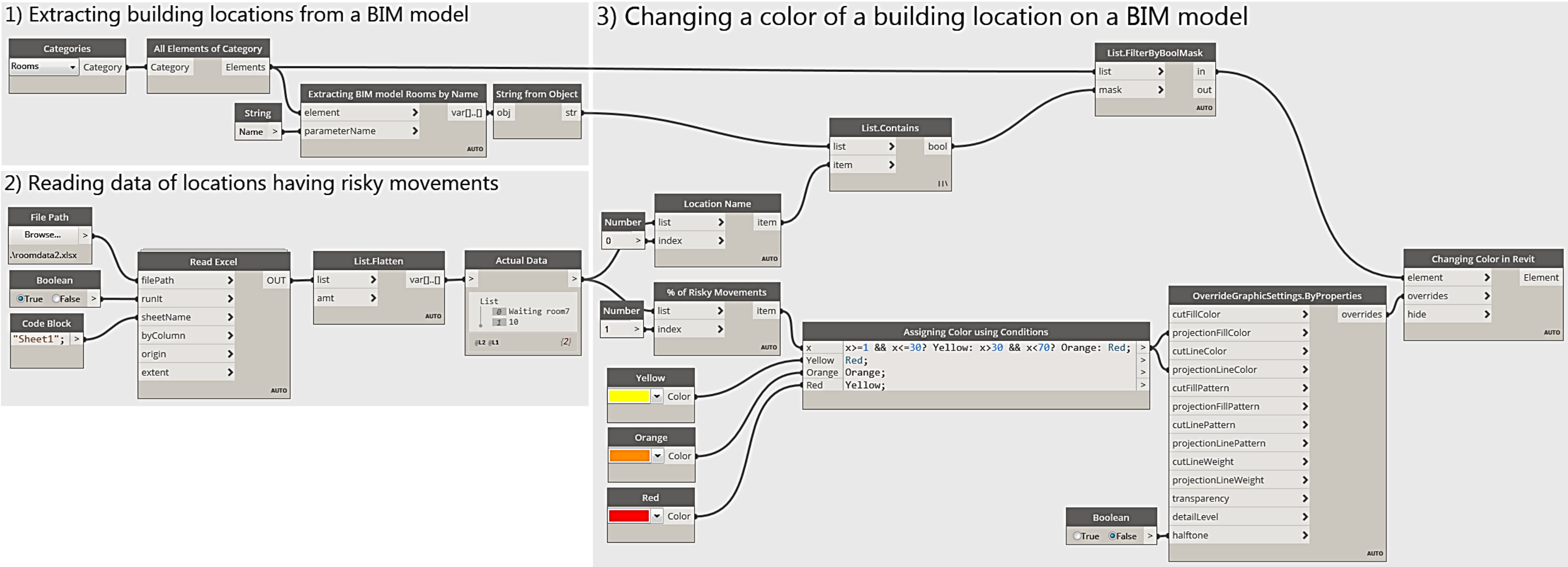

Fig. 5-11 A Dynamo graph for visualizing different colors against risky movements on a BIM model

Visualizing different types of risky movements on a building model in Revit software - Red for 70% or more, Orange for greater than 30% and less than 70%, and a Yellow for less than or equal to 30% occurrences of risky locations in a trajectory

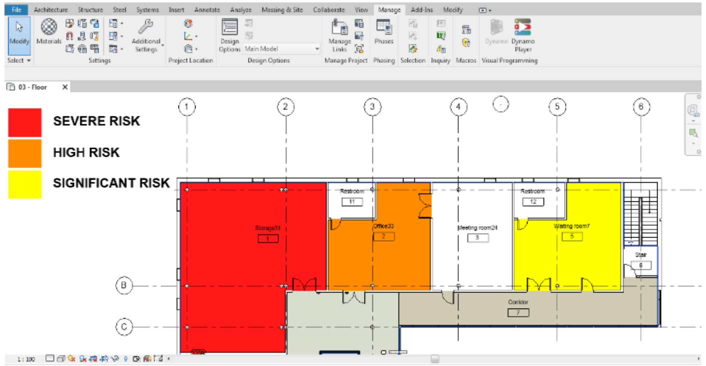

Fig. 5-12 Visualizing different types of risky movements on a building model in Revit software

‘OverrideGraphicSettings.ByProperties’ node. This node changes the color of the Revit element (room in our case) in the active view irrespective of the current element display properties. The descriptions and the functionalities of the nodes which were used for constructing our Dynamo graph (see Fig. 5.11) can be found in (Autodesk Dynamo, 2019).

5.3 Discussion

The experimentation for understanding the user trajectories described is done using the batch processing techniques. At first, a trajectory dataset of building users is acquired using BLE beacons and then pre-processing is performed in an offline mode. However, to achieve the real-time insights of worker movements on construction sites, stream processing methods should be employed for pre-processing the worker trajectories in an online mode. Whereas, a process of the HMMs` computations should remain a batch process as their dependency on the previous observations. After collecting the required dataset of trajectories, workers` stay regions are identified with their stay duration. The reason for extracting such locations in trajectories is to identify the ROIs and their associated POIs. The ROIs are the wider areas of the building (e.g. work-zone237, etc.) which include multiple geographical POIs (e.g. outdoor pathway, storage room, office1, etc.) labeled as ‘rooms’ in our model. The idea of splitting a building floor into a set of ROIs is taken from the Pradhananga and Teizer, (2013) research. As construction sites are divided into different zones such as material zone, dumping zone, loading zone, work zone, etc. by the building supervisors for easy and well-organized site monitoring and management. The developed WoTAS system has identified several stay regions in users` trajectories that correspond to different ROIs. For mapping, the processed trajectories with the ROIs, spatial joins (Z. Yan, 2011) are performed as discussed in Chapter 4. The extraction of the ROIs mapped to a set of users` trajectories is shown in Fig. 5.13.

user	location	concept	start	end	duration
Maintenance 1	Storage-zone-300	stride:Storage	2018-01-02T09:05:00	2018-01-02T12:00:00	175
User 1	Work-zone-237	stride:SafeWorkzone	2018-01-02T09:06:00	2018-01-02T09:59:00	53
User 1	Work-zone-237	stride:DangerousWorkzone	2018-01-02T10:01:00	2018-01-02T11:15:00	74
User 1	Storage-zone-300	stride:Storage	2018-01-02T11:21:00	2018-01-02T12:00:00	39

Fig. 5-13 ROIs of a user

After labeling the ROIs and the POIs, an entire set of trajectory points belonging to an ROI is further analyzed using the HMMs for identifying unsafe user movements having long steps and many turnings per minute within different POIs. Though, the insights can be extracted on the data sampled per second for understanding the worker movements more thoroughly. The same process is repeated for each identified ROI to have a complete picture of different movements occurring inside all the critical building locations. The purpose here is to understand the movements within the trajectories with respect to identified POIs corresponding to different building zones (ROIs). As shown in Fig. 5.10, the values of the hidden states are variations in the movements observed during different times in a day within an ROI named ‘work-zone237’. There are unsafe movements identified having long steps or many turnings as the 3rd state in Fig.

5.10. The hidden states` values correspond to different POIs are used for plotting the building model for visualizing the most probable user movements within the identified ROI as shown in Fig. 5.12. For further tracking, the unsafe movements, the corresponding speed and turning angle values associated with such movements can be traced from the time series plots as shown in Fig. 5.5 – 5.8 for deeper analysis.

Visualizing the categorization of movements using step lengths and turning angles using a BIM software will help H&S managers in monitoring critical building locations for identifying potentially unsafe incidents in the buildings or on the construction sites and enabling them in taking quick actions in the event of unsafe worker movement detection. In addition, the developed system will increase the spatial awareness level of sites by letting building supervisors know the potentially unsafe locations so that safe distances can be maintained between the worker and the construction machinery involving high speed or many turnings. In this way, the occurrences of accidents resulting from the excessive proximity between the workers and the machinery can be reduced significantly. Moreover, if the construction machinery is not being operated in the safe speed limits on a construction site, our BIM-based movement visualizations can help the H&S managers in identifying the workers who are not complying the safety regulations for operating the machinery and eventually the need of additional training for the workers can be determined accordingly. Furthermore, understanding the categorized movements of the workers and the machinery will act as a pro-active measure during the planning of future construction operations by the building supervisors to avoid unsafe situations.

Chapter 6 - Detecting intrusions for safety management

Outline

This chapter is based on the 2nd case-study to use OBiDE framework by developing a prototype system named 'VIDEWS' (Visualizing Intrusions in Dynamic Environments) to identify near-miss incidents such as intrusions on sites aiming to reduce the construction site fatalities. The contributions of this chapter are:

(i) ***Intrusion detection subsystem:*** After constructing the trajectory pre-processing (i.e. data cleaning) and semantic enrichment (using an OSM file of a building) system as described in Chapter 4, an entire set of trajectories is further analyzed for detecting the intrusions in a building.

(ii) ***BIM-based visualization subsystem:*** The output of the intrusion detection system (a triplestore) is linked using a BIM model for visualizing the building locations where the intrusions have occurred.

It should be noted that the fundamental concepts related to data acquisition, semantic enrichment and historicization of occupant trajectories are the same for both proposed applications of the OBiDE framework. To avoid the repetition of concepts, this Chapter will only present the extraction of worker intrusions for a safety management application. Whereas, to study the underlying methods behind this application, see Chapter 3 to 5.

6.1 Scenario and the Prototype System

For monitoring the worker movements to identify intrusions in buildings, an OBiDE Framework, designed for dynamic environments, is applied (see Fig. 6.1) which is. In this case,

1) 'driver' is monitoring movements of occupants,
2) 'need' is achieving safety management in a building by identifying unsafe movements,
3) 'action' is tracking movements of occupants using their spatio-temporal trajectories,
4) 'system' corresponds to Bluetooth Low Energy (BLE) beacons for sensor data acquisition deployed in a building, and
5) 'states' are; 1) safe movements and 2) unsafe movements (intrusions).

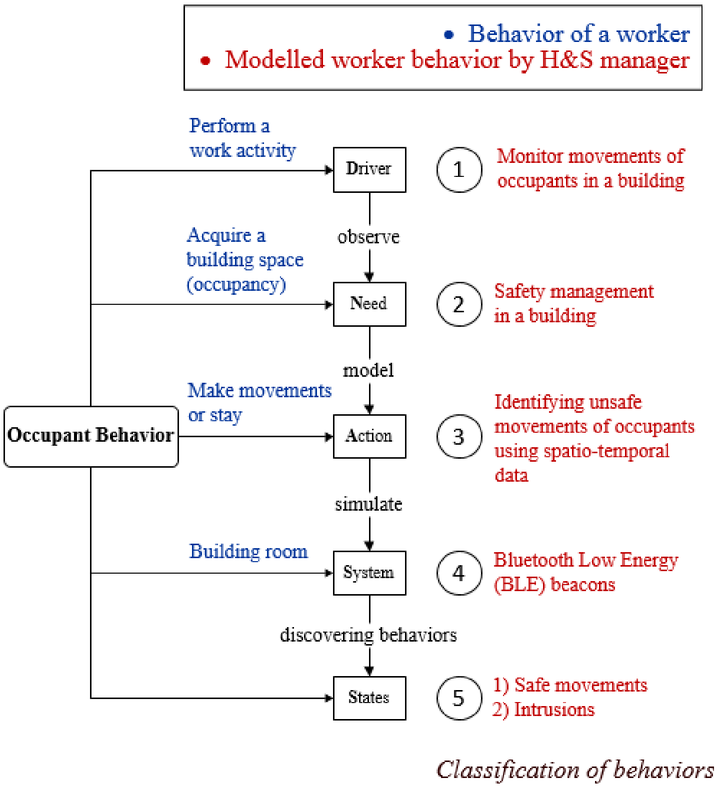

Fig. 6-1 Description of a scenario using OBiDE framework

After describing the required behaviors, a scenario is initially created to detect intrusions on the construction sites by considering the roles of building supervisors and H&S managers. Based on the scenario (see Fig. 6.2), the prototype system named 'VIDEWS' is developed. The working of the scenario initiates as the Graphical User Interface (GUI) for extracting the information of detected intrusions is invoked from the BIM software. The system interactions which were involved in creating this GUI functional are stated below;

1. For collecting the worker movements in a dynamic building environment, the BLE beacons are mounted on different locations in buildings.

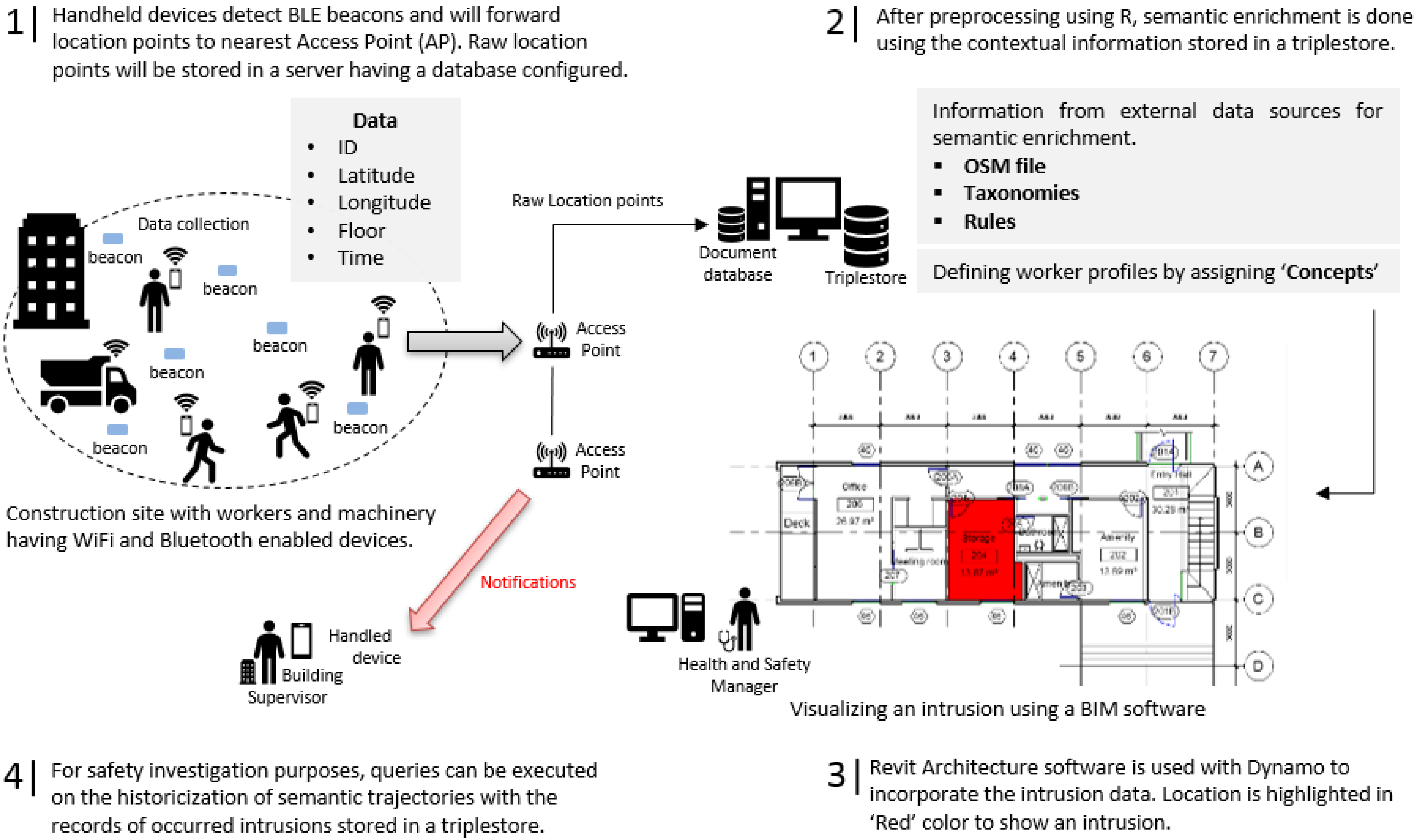

Fig. 6-2 Prototype system scenario for identifying intrusions on a construction site

The handheld devices of workers capture three strongest detectable signals of neighbouring BLE beacons, estimate the locations in the form of spatio-temporal points using the beacon deployment map information and forward them to the nearest Access Points (APs) for storing them in a centralized database.

2. The spatio-temporal points of the workers are enriched with their corresponding building locations using the geographical information extracted from the OSM files of buildings and later stored in an ontology-based STriDE data model (i.e. a triplestore) with the information of worker profiles for historicization.
3. Using the information stored in a triplestore, the intrusions are detected. The output of a triplestore is linked with the BIM software for visualizing the building locations where the intrusions have occurred.
4. All the information on intrusions with the building evolution will be stored in a triple store for future investigation purposes and will help in conducting cause and effect analysis on the occurrences of near-misses on sites.

The scenario presented above is created ideally for the construction sites where the locations evolve in terms of location, size, properties, and relationships with the environment. However, the functionality of the scenario is validated in two already constructed educational buildings where the locations only evolve in terms of contextual information (i.e. alphanumeric properties). BLE beacons are chosen for spatio-temporal data collection of building users because of their portability, easy to deploy and low-cost solution (Paek et al., 2016). The geographical information extracted from the OSM files of buildings is used for constructing the beacons` deployment plan to cover all the possible building locations for data collection. Real-time user movements in two buildings are collected in a batch mode (Zheng, 2015) for a period of 2-weeks. Later, the user movements are enriched with their corresponding building locations and user profiling is achieved. Lastly, the detected intrusions are visualized using BIM software. The BIM files of buildings where the movement data is collected do not exist. To demonstrate the proof-of-concept integration of systems and the working of Revit API, a sample BIM model is used for linking it with the information of intrusions extracted using a triplestore. For integrating the information of intrusions with corresponding building locations in a BIM model, unique room identifications are used and made identical in BIM as well as in the STriDE model. The details of the subsystems of a developed prototype are mentioned below;

6.1.1 Detecting intrusions

After generating the semantic trajectories (see Fig. 6.3) using the stored information in the STriDE model users` trajectories are represented as a series of timeslices (see Fig. 6.4) for detecting intrusions in a building. In our model, as already described above that ‘concepts’ are defined for tagging building locations with spatio-temporal trajectories. These concepts are defined using SKOS vocabulary and stored in the ‘concept scheme’. A concept scheme is a collection of different concepts, described as a hierarchy, that correspond to different building locations and are stored as triples in the RDF file. The purpose of defining a hierarchy is to be more precise about the tagging of a building location. For instance, for recognizing a specific construction site region, we can move towards more exactness to call it as a hazardous zone or a work zone. For achieving user profiling, concepts are used as shown in Fig. 4.2. Here, a profile is defined as a set of different concepts that a user has access to. For identifying the intrusions, the tagged concepts with the user timeslices are mapped with the allocated concepts according to the user profiles. In this way, unauthorized locations are recognized, and alerts are generated to alert the user and the building supervisors. A triple store query is used for extracting an intrusion from a list of users` timeslices. Fig. 6.5 is the screenshot of the identified intrusion.

traj	userName	location
stride:TrajOfWorker2-1	Worker 2	Outdoor pathway
stride:TrajOfWorker2-2	Worker 2	Storage room
stride:TrajOfWorker1-1	Worker 1	Outdoor pathway
stride:TrajOfWorker1-2	Worker 1	Corridor of floor 0
stride:TrajOfWorker1-3	Worker 1	Office 1
stride:TrajOfWorker1-4	Worker 1	Corridor of floor 0
stride:TrajOfWorker1-5	Worker 1	Outdoor pathway
stride:TrajOfWorker1-6	Worker 1	Storage room

Fig. 6-4 Trajectories database having timeslices of two users

s	p	o
stride:TrajOfWorker1-3	rdf:type	stride:TimeSlice
stride:TrajOfWorker1-3	stride:hasStartDate	2018-01-02T09:06:00
stride:TrajOfWorker1-3	stride:hasEntity	stride:TrajOfWorker1
stride:TrajOfWorker1-3	stride:hasEndDate	2018-01-02T09:30:00
stride:TrajOfWorker1-3	stride:isTrajectoryOf	stride:Worker1
stride:TrajOfWorker1-3	stride:hasLocation	stride:W5

Fig. 6-3 Timeslice description of a Worker1 trajectory

userName	room	roomLabel	start	end
Worker 1	stride:W1002	Storage room	2018-01-02T09:36:00	9999-12-31T23:59:59

Fig. 6-5 Detecting an intrusion from trajectory`s timeslices

6.1.2 Visualizing intrusions using BIM

The building in which the experimental set up is done for collecting spatio-temporal trajectories of workers, its BIM model doesn't exist. Eventually, a Revit architectural sample file is used and modified using the information extracted using an OSM file of a building for representing a proof-of-concept integration of systems. Sensors (beacons in our case) are typically not part of a Revit model. For incorporating the sensor data (semantically-enriched trajectories) into a BIM model, Dynamo is used. Dynamo works alongside a Revit software requiring an active Revit document having a BIM model and executes a simple execution structure that is; input, process, and output (Autodesk Dynamo, 2019). It means that firstly the relevant building geometry is taken from Revit into Dynamo, then desired processes are performed using pre-packaged nodes, and finally, the Dynamo outputs the results back into the Revit document. Constructing a Dynamo script (also called a graph) consists of 'nodes' and 'wires'. Nodes are the blocks of code that aim to execute a single discrete function. Whereas, wires are used to connect nodes` outputs to nodes` inputs for building the data flow from left to right (Autodesk Dynamo, 2019). Dynamo graph constructed for the 'VIDEWS' system consists of four key steps (see Fig. 6.6), which are;

1. All the building locations which are tagged as 'rooms' in the BIM model are extracted by defining the category of elements as 'rooms' in the Dynamo. In our case, the parameter 'room name' will act as a unique identifier for each space for visualizing the problematic room in the Revit model where an intrusion has occurred. The naming convention of Revit rooms is set according to the tagging of semantic locations as described in the STriDE model.

2. After extracting the list carrying room names of the building, the details of occurrence of intrusions i.e. intruder ID, time of occurrence of intrusion and room name where the intrusions are detected are imported into Dynamo. This is achieved by calling an API to output the result from the STriDE model automatically in the form of an Excel sheet containing the details of occurred intrusions.

3. On receiving room`s data from an active Revit document and intrusions` data taken from the STriDE model, the room location where an intrusion has occurred is highlighted in 'Red' color in a BIM model using the mapping of room names in both lists (see Fig. 6.7). The change in the color of the Revit room is done using the 'Element: OverrideColorInView' node. This node changes the color of the Revit element (room in our case) in the active view irrespective of the current element display properties.

4. As soon as the problematic room is highlighted in the Revit, this process invokes another user window to ask for details of the intrusion that occurred in a specific room location. After inputting the desired room name, time of occurrence of the intrusion in the particular room and its corresponding intruder ID will be displayed to the user in the Revit software.

6.2 Discussion

BBS training has a potential for reducing near-miss incidents such as intrusions on the construction sites by educating the workers and increasing their safety awareness about the dynamic nature of construction job environments (Fang et al., 2018; Heng et al., 2016). However, BBS training completely depends on the manual procedures executed by the experienced building supervisors. Also, it lacks timely feedback, and an assessment of the occurrence of intrusions for further construction operations. To overcome this limitation, intrusion detection systems based on different sensor-based technologies are developed in the existing literature. However, these systems cannot recognize intrusions in dynamic environments where the building locations evolve (Fang et al., 2018; Heng et al., 2016; Teizer et al., 2015). This opens many data

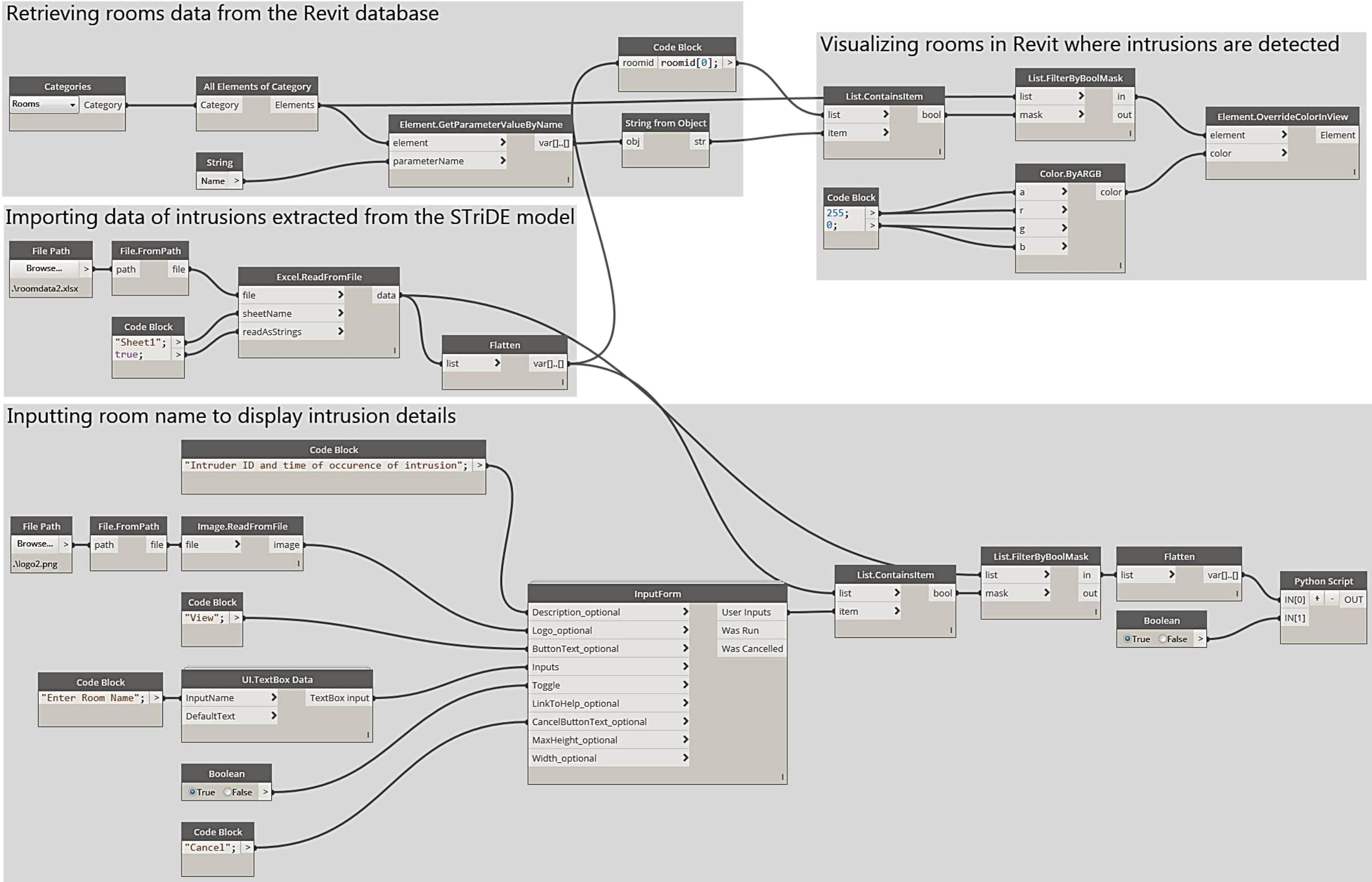

Fig. 6-6 Graph in Dynamo for visualizing intrusions

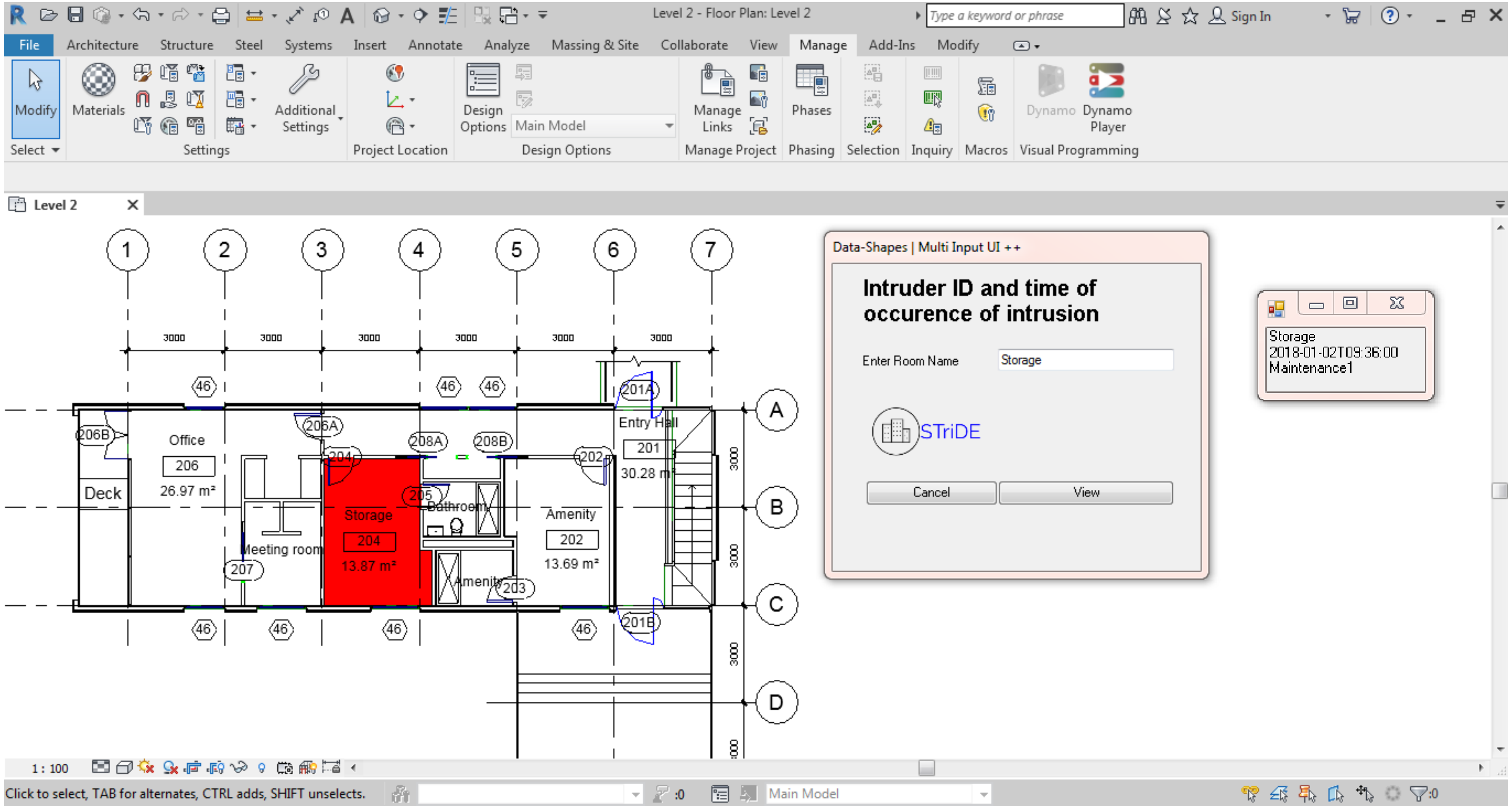

Fig. 6-7 Visualizing intrusion information in Autodesk Revit software

management challenges for keeping track of the changes in the contextual and spatial information associated with the locations which evolve for identifying intrusions from the perspective of the building or site environment. The main contribution of this study was the development of an intrusion detection system based on a data model which can track the building spatial and contextual evolutions using spatial-temporal and filiation relationships among building entities (users and locations).

The developed system is validated using a real-life IPS feed of the building users extracted from BLE beacons. BLE beacons originally provide proximity-based services and coarse-grained location positioning for indoor application scenarios (Paek et al., 2016). The developed system holds the information at room level that is collected per second using beacons for identifying each building user using its unique identification to execute an intrusion detection system. Later, the movements are processed and enriched with their associated building locations using an OSM file. An OSM file has provided the geographical coordinates of the buildings. This geographical information is used for deploying the beacons in buildings and defining the building locations in a data model. However, a BIM file can also be used for beacon planning, deployment, and retrieval of building infrastructural information. But the system is implemented using an OSM file as BIM files of the buildings where the experimentations are conducted do not exist. However, the output of the developed system is integrated with the BIM software to visualize the building locations where the intrusions have occurred. In the literature (Riaz et al., 2014), there exists many 3-dimensional Computer-Aided-Design (CAD) software for generating the building visualizations for different built environment applications. Though, a BIM approach is chosen as it enhances the inter-organizational collaboration of information among project members because it is based on the Industry Foundation Classes (IFC) standard that is universal and supports easy and fast information exchanges (Volk et al., 2014).

Chapter 7 Discussion and Conclusions

Outline

This chapter first presents the overall synthesis of the work done for developing the OBiDE framework. Second, it describes how the developed safety management systems can contribute towards reducing construction accidents. Third, it presents a brief cost-benefit analysis of developed systems. Lastly, it describes the conclusions of this research along with some future directions.

7.1 Contribution towards Generalization of Occupant Behavioral Framework

The main research question which guided this research was, how to establish a high-level understanding of the human behaviors in a dynamic environment using a generalized framework that should be constructed using the fundamental components of the state-of-the-art systems for human behavioral understanding?

The research question was motivated by the fact that human behaviors evolve as well as the buildings. Responding to human actions (movements) when they have been executed allows analysts to examine their complete effects in the environment. But the outcomes of their behaviors may create serious problems for other humans and the environment if the evolution of a building is not considered for modeling and analysis of occupant-building interactions. Existing systems (as described in Chapter 2) in the literature are built by capturing the stochastic and reactive nature of occupant behaviors for modeling and building analysis. Ultimately, these systems contribute to enhancing the understanding of occupant behaviors by increasing the occupancy resolution (i.e. inferring different occupant activities using their spatio-temporal behaviors) for different building management applications. However, the existing systems for occupant behavior modeling do not incorporate the information of complex dynamic environments where the building objects (occupants and building locations) evolve. Over time, the functionalities of the locations in a building often change (i.e. change in semantics or a context). For example, a room named 'inventory' in a building is now an 'office' having different functionality. Likewise, due to the placement of certain inventory in the specific area of a building, the floor area of a building (a set of rooms and a corridor) became a 'restricted area'. Also, new walls or infrastructure support may be added to a building. This will result in a change in the dimensions (i.e. geometry) of building locations (called as spatial changes). The change in the semantics of building locations occurs often in constructed facilities whereas, the spatial changes take place rarely. Such changes need to be incorporated in occupant behavior modeling as a change in the purpose or a position of building locations will result in different behaviors of occupants which ultimately represent different occupant activities. The updated spatial and contextual information about the building locations along with the previous information will contribute to an improved understanding of occupant behaviors concerning the changes occurred in the building environment. To address these requirements of dynamic building environments that contain evolving building objects, an integrated framework named 'OBIDE' is proposed (see Chapter 3).

There are six different processes which were involved in constructing the OBiDE framework which are; 1) defining human behaviors using DNAS ontology, 2) capturing behaviors of occupants (more specifically, occupants` movements and presence, the prerequisites for any kind of human behavior understanding), 3) preprocessing (filtering, segmentation, and stay points detection) the occupant behaviors, 4) semantic enrichment using the building information extracted from an OSM file and historicizing the behaviors, 5) modeling the behaviors by incorporating the secondary data sources according to the use-case applications and categorizing the behaviors based on their extracted states, and finally 6) visualizing the classified occupant behaviors using BIM for building managers.

7.2 Contribution towards Reducing the Construction Accidents

The best example of a dynamic environment to study different occupant behaviors in the context of an evolving building environment using the proposed OBiDE framework is the construction site, where spatial as well as semantic changes occur often. Considering the construction site scenario, OBiDE framework is applied for analyzing the worker behaviors using the historicized information of worker movements mapped with the relevant building contextual information for two different safety management applications ('WoTAS' - Worker Trajectory Analysis System and 'VIDEWS' - Visualizing Intrusions in Dynamic Environments). These applications are discussed in Chapters 5 and 6. For both the safety applications, initially, beacon-based data collection and trajectory pre-processing subsystem are built for extracting multifaceted trajectory characteristics. Second, the collected trajectory points are mapped with their corresponding Region of Interests (ROIs) and their associated Position of Interests (POIs) locations. The ROIs are the wider areas of the building (e.g. work-zone237, etc.) which include multiple geographical POIs (e.g. outdoor pathway, storage room, office1, etc.) labeled as 'rooms' in the model.

For the 1st construction safety application (i.e. 'WoTAS') that is built for worker movement analysis, the Hidden Markov Model (HMM) is used which is one of the probabilistic approaches for describing the object behavior in time. Using the HMMs, trajectories are analyzed by categorizing the worker movements into three different states. In the end, the output of the Viterbi algorithm is visualized using a BIM model for identifying the high-risk locations involving sharp worker movements and rotations. For the 2nd construction safety application (i.e. 'VIDEWS'), that is built for detecting intrusions in buildings. With the help of occupant profiling that is achieved in a data model, intrusions are extracted and later visualized on a BIM model.

Existing literature highlights that majority of the construction site accidents occur because of any inadequacy in the a) safety training (to improve the safety consciousness of the worker by training), b) safety planning (identifying the Job Hazard Areas (JHAs) by team meetings), and c) construction site planning (identifying and recording worker violations by maintaining the checklists) (Teizer et al., 2013; Guo et al., 2017). The safety management applications ('WoTAS' and 'VIDEWS') presented in this thesis aims at contributing towards reducing the chances of occurrence of all three causes as mentioned in Table 7.1.

Table 7-1 Developed system contributions towards reducing the accidents

	Accident cause	Contribution
1	**Inadequate worker training**	The developed systems ('WoTAS' and 'VIDEWS') monitor the worker movements using their spatio-temporal trajectories. However, the extracted movement state i.e. 'long steps and many turnings' or an intrusion representing an unsafe behavior of a worker does not necessarily lead to an accident. However, it will help in identifying the high-risk activities of the workers (Teizer and Cheng, 2015).

		For instance, the high-risk activities involving fast movements and sharp rotations or detecting intrusions will increase the situational awareness of the H&S managers in identifying the needs for more safety training programs required by the certain workers who are violating the speed limits on sites, not operating the machinery in safe working mode or accessing the building areas where they are not authorized to work in (Teizer et al., 2013). Moreover, the workers who are staying for extended durations than the required time for work activities within the certain regions on sites which contain hazardous substances such as dust, fumes or explosive materials can be dangerous to worker`s health and safety (Li and Guldenmund, 2018; Dong et al., 2017). The workers violating the safety regulations can be identified using our developed systems by extracting the stay regions from trajectories after setting the required time and distance thresholds (see Fig. 4.12).
2	**Incomplete site planning**	Failure in identifying the high-risk regions (JHAs) on a construction site, where potential work hazards lie is a major cause of accidents (Guo et al., 2017). Without technological and visualization support, it is unrealistic to identify the high-risk regions on the whole area of a site simultaneously because of its huge size and dynamicity of the environment and eventually numerous JHAs go undetected (Guo et al., 2017). The proposed BIM-based developed systems cater to this issue by integrating the information of the worker movement behaviors and the on-site environment (i.e. OSM file). Certain JHAs concerning; 1) the high dense site regions (i.e. extracting the ROIs) which require special space demands, 2) maintaining the safe distances between the operating machinery and the workers for preventing the spatial collisions (Teizer et al., 2013; Guo et al., 2017), and 3) visualizing the worker movements in conjunction with the locations of the risky places such as holes and edges (by viewing building elements` data from a 3D building model) which may lead to falling accidents Guo et al., 2017) can be addressed by proposed BIM-based systems.

		For more information on how this research addresses the aforementioned JHAs` concerns, see the article by Arslan et al. (2018).
3	**Invalid site monitoring**	A process of monitoring the construction site encompasses five major steps (Golparvar-Fard et al., 2009) which are; 1) collecting the site data (workers and the building environment), 2) converting the captured data into usable information, 3) identifying the current situation on a site from the acquired information, 4) reporting the H&S managers about the situation on time so that actions can be taken, and 5) recording the actions taken. The developed systems (‘WoTAS’ and ‘VIDEWS’) attempts to improve the process of site monitoring (see Chapter 5 and 6). It can capture the worker movements using low-cost BLE beacons as well as the evolving site environment. The process of semantic enrichment employed for the systems ensures that the captured data is enriched with the relevant contextual information. The extraction of the ROIs and the POIs will make the H&S managers informed about the current site situation by identifying the important regions of a site where most of the workers are staying and relevant site resources can be managed accordingly. Also, the developed systems will act as a feedback system that will alert the H&S managers in case of high-risk activities` detection using a BIM-based visualization. Moreover, the developed system supports the historicization of the semantic trajectories using a triplestore that is useful for querying the workers` movements and analyzing the impact of dynamic locations on the worker trajectories for safety analysis in the future.

7.3 Cost-Benefit Analysis of Developed Safety Management Systems

The developed ‘WoTAS’ application contains five subsystems which are; 1) data collection subsystem, 2) pre-processing subsystem, 3) semantic enrichment subsystem, 4) HMM-based movement analytics subsystem, and 5) BIM-based visualization subsystem. Whereas, ‘VIDEWS’ prototype system consists of the same subsystems as of ‘WoTAS’ except the HMM-based movement analytics subsystem. The developed systems present several benefits with minimal costs of deployment and maintainability in buildings as discussed in Table 7.2.

Table 7-2 Cost-benefit analysis of ‘WoTAS’ and ‘VIDEWS’

No.	Subsystem	Technology used	Description with benefits	Cost
1	**Data collection**	BLE beacons	Low battery powered 200 beacons captured 8,426 trajectory points in a building of 11 users in a time period of 2 weeks. Out of 110 locations in collected trajectories, 107 were correctly tagged. The tagging accuracy was high for our indoor scenario. BLE beacons didn’t had any issue with the Android-based smartphones as the beacons were easily detectable. An Android platform is chosen for performing geo-localization and data transmission to a server using WiFi access points as it is one of the most widely used operating systems in hand-held devices. One of the main benefits of using a beacon-based technology is its portabilility. Beacons can be placed anywhere on construction sites based on the requirements because of their small size and are completely reusable (Gomez-de-Gabriel et al., 2019). In addition, beacons are very simple to program, deploy and integrate with wearable devices (Kontakt, 2019; Ciabattoni et al., 2019).	The cost of BLE beacons ranges from 10 to 20 euros per unit (Gomez-de-Gabriel et al., 2019) and is totally dependent on the size of the area which needs to be monitored.
2	**Pre-processing**	R Studio with document database (MongoDB)	After data collection, R studio is used with a cloud-based document database for data cleaning. The benefit of a pre-processing subsystem is to keep logs of worker movements (cleaned trajectories) so that the stored information can be used for safety analysis. For example; extraction of the stay regions of the workers in buildings, segmenting worker trajectories based on speed values into run and walk episodes, finding the number of workers inside a pre-defined	R studio and MongoDB are open-sourced softwares.

			area (defining explicitly) to avoid site congestion and calculating the safe distances between different workers to avoid transportation accidents using Haversine distance formula. For more information on how this research addresses the aforementioned JHAs` concerns, see the article by Arslan et al. (2018).	
3	**Semantic enrichment**	Triplestore (Stardog)	Semantic enrichment subsystem has label each trajectory point by tagging it with the relevant contextual information. For creating a knowledge base, a triplestore is used to hold the semantically-enriched trajectory data. This step has provided us a way to perform historicization of the semantic trajectories which is used for movement-based trajectory analysis of workers using the context of an evolving building.	The licensed version of a triplestore is used. However, open-sourced triplestores are also available.
4	**Movement analysis**	R Studio	For categorizing the worker trajectories into different movement states for each stay region, the HMMs are used. Categorizing movements in trajectories will help the H&S managers to monitor critical building locations for identifying accident-prone scenarios, and enabling them to take quick actions in case of detection of unsafe worker movements. In addition, this understanding will act as a pro-active measure for preventing fatalities and the planning of future construction operations can be readjusted by the building supervisors to avoid unsafe situations.	R studio is open-sourced.

5	**BIM-based visualization**	Revit with Dynamo	For generating trajectory visualizations using a BIM-based Revit Architecture software, Dynamo (a visual programming software) is chosen. Dynamo eliminated the need for exhaustive programming which sometimes is not convenient for the safety managers to perform for achieving desired visualizations (Amann et al., 2018). Safety managers can use Dynamo as per their needs for creating different variants of developed visualizations containing the information without acquiring the advanced programming skills.	Autodesk Revit software is used with a student license. Whereas, Dynamo is open-sourced.

7.3 Conclusions and Future Works

Occupant behavior is a crucial and often overlooked factor in attaining building performance and management goals such as monitoring occupants for safety or reducing the consumption of building resources (e.g. energy, materials, etc.). Though advanced smart building technologies integrate different building subsystems to optimize the usage of the building resources as well as ensuring safety in buildings. But occupant behaviors (e.g. actions) if not incorporated in modeling can significantly impact the overall performance of a building. The buildings are subject to constant change, where the locations evolve in terms of geometry and contextual information during a building lifecycle. This makes the data collection of occupants as well as the building infrastructural changes a continuous process for the exploration of occupant behaviors using a building context. This study has presented an integrated OBiDE framework to maintain the dynamicity of a building environment in a data model and fusing the occupants' movements and presence data (the prerequisites for any type of occupant behaviors' understanding) with the most updated building information.

Managing the safety of workers in a building using 'WoTAS' - Worker Trajectory Analysis System and 'VIDEWS' - Visualizing Intrusions in Dynamic Environments for understanding their mobility-related behaviors are two of the possible use cases of the OBiDE framework. The focus while implementing the OBiDE framework was not to perform tests for finding the best modeling technique (i.e. HMM for this research) for extracting unsafe occupant behaviors but to show the integration of systems for fusing trajectory data with the evolving building information to identify unsafe worker movements. However, comparing the extracted patterns from different probabilistic models and evaluating their results based on the real-time site studies can be done as a future work of this research which will help in selecting the most optimal probabilistic model for extracting the behavioral states and classifying the occupant behaviors.

One of the limitations of developed systems is that the building information is extracted using an OSM file for feeding into a data model for achieving the historicization of buildings. However, BIM model (i.e. an industrial standardized method for maintaining building lifecycle data) should be utilized for all the building-related processes which include sensor placements in buildings and spatial information mapping with the occupant locations extracted from BLE beacons. Moreover, the developed system should be made synchronized with the updates occurring in the BIM model in real-time for maintaining the historicization of building information by tracking the spatial and contextual evolutions in dynamic construction environments. Furthermore, at present, the prototype system is implemented in already constructed buildings. Further work is required to empirically test the functionality of the developed system on real construction sites to understand the unsafe behaviors of workers to reduce safety-related accidents in buildings.

The OBiDE framework which tracks spatio-temporal data of occupant generated in a dynamic environment establishes a basic reference model that can not only be used for presented safety management applications but also for space utilization analysis and energy performance simulations where the decision-making processes are performed based on the occupant movements across different locations in buildings. However, additional behavioral occupant datasets should be collected that could allow an improved understanding of a dynamic building environment for enhancing the building overall performance by increasing the occupants' productivity, environmental quality in a building, and functionality of building spaces.

Furthermore, for the future direction of this research, it needs to be mentioned that currently this research is solely based on computing concepts that were utilized based on the construction site use case applications. However, a multidisciplinary approach is needed for developing the behavioral models for better understanding the occupant behaviors. And this can be achieved by combining the knowledge of other than movement-related behavioral, social and psychological factors that contribute to human actions (behaviors) over time.

List of Publications

Journal Publications

1) **Understanding Occupant Behaviors in Dynamic Environments using OBiDE Framework**
By **M. Arslan**, C. Cruz, D. Ginhac, Building and Environment 166, Dec. 2019, p.106412. https://doi.org/10.1016/j.buildenv.2019.106412 (SCImago journal rank: Q1, Impact Factor: 4.8)

2) **Visualizing Intrusions in Dynamic Building Environments for Worker Safety**
By **M. Arslan**, C. Cruz, D. Ginhac, Safety Science 120, Dec. 2019, pg. 428-446. https://doi.org/10.1016/j.ssci.2019.07.020 (SCImago journal rank: Q1, Impact Factor: 3.6)

3) **Semantic Trajectory Insights for Worker Safety in Dynamic Environments**
By **M. Arslan**, C. Cruz, D. Ginhac, Automation in Construction 106, Oct. 2019, p.102854.https://doi.org/10.1016/j.autcon.2019.102854 (SCImago journal rank: Q1, Impact Factor: 4.3)

4) **Semantic Enrichment of Spatio-temporal Trajectories for Worker Safety on Construction Sites**
By **M. Arslan**, C. Cruz, D. Ginhac, Personal and Ubiquitous Computing, Jan. 2019. pp. 1-16. https://doi.org/10.1007/s00779-018-01199-5 (SCImago journal rank: Q1, Impact Factor: 1.7)

5) **Spatio-temporal dataset of building occupants**
By **M. Arslan**, C. Cruz, D. Ginhac, Data in Brief 27, Oct. 2019, pp. 1-7, p.104598. https://doi.org/10.1016/j.dib.2019.104598 (SCImago journal rank: Q1, CiteScore: 0.93)

6) **Spatio-temporal Analysis of Trajectories for Safer Construction Sites**
By **M. Arslan**, Christophe Cruz, Ana-Maria Roxin, Dominique Ginhac, Smart and Sustainable Built Environment, Vol. 7 Issue: 1, pp.80-100, March 2018. https://doi.org/10.1108/SASBE-10-2017-0047 (SCImago journal rank: Q2, CiteScore: 1.04)

Conference Publications

1) **DNAS-STriDE Framework for Human Behavior Modeling in Dynamic Environments**
By **M. Arslan**, C. Cruz, D. Ginhac, 19th international conference on Computational Science – ICCS 2019, 2-14 June, 2019, Faro, Portugal, Lecture Notes in Computer Science, vol 11540. Springer, Cham, pp 787-793. https://doi.org/10.1007/978-3-030-22750-0_79

2) **Identifying Intrusions in Dynamic Environments using Semantic Trajectories and BIM for Worker Safety**
By **M. Arslan**, C. Cruz, D. Ginhac, 4th international congress on information and communication technology, February 25 - 26, 2019, London, United Kingdom. Springer AISC**. ISBN Number - 2194-5357

3) **Movement Behavior Analysis of Workers using Spatio-temporal Trajectories for Safety Management**
By **M. Arslan**, C. Cruz, D. Ginhac, MARAMI 2019, Modèles & Analyse des Réseaux: Approches Mathématiques & Informatiques, The 10th Conference on Network Modeling and Analysis, November 06 - 08, 2019 Dijon, France.

.

4) **Exploiting Semantic Trajectories using HMMs and BIM for Worker Safety in Dynamic Environments**
By **M. Arslan**, C. Cruz, D. Ginhac, 5th International Conference on Computational Science and Computational Intelligence in 'Smart Cities & Smart Mobility', IEEE, December 13-15, 2018, Las Vegas, Nevada 89119, USA.
https://americancse.org/events/csci2018/schedules/CSCI_2018_pdf

5) **Understanding Worker Mobility within the Stay Locations using HMMs on Semantic Trajectories**
By **M. Arslan**, C. Cruz, D. Ginhac, 14th IEEE International Conference on Emerging Technologies, Nov 21-22, 2018, Pakistan. https://doi.org/10.1109/ICET.2018.8603666

6) **Semantic Enrichment of Spatio-temporal Trajectories for Worker Safety on Construction Sites**
By **M. Arslan**, C. Cruz, D. Ginhac, 9th International Conference on Ambient Systems, Networks and Technologies (ANT 2018), Porto, Portugal, Procedia computer science (Elsevier) 130, pp.271-278. https://doi.org/10.1016/j.procs.2018.04.039

7) **Using Spatio-temporal Trajectories to Monitor Construction Sites for Safety Management**
By **M. Arslan**, C. Cruz, A. Roxin, D. Ginhac, ICIME 2017, October 9–11, 2017, ACM, pp. 211-216, Barcelona, Spain. https://doi.org/10.1145/3149572.3149600

8) **Big data Applications for Disaster Management**
By **M. Arslan**, A. Roxin, C. Cruz, D. Ginhac, BigCVEn-13th International Conference on Signal image processing technology and internet-based systems, IEEE, December 4-7, 2017, pp.1-6, Jaipur, India. https://ieeexplore.ieee.org/document/8334773

Published Dataset

1) **Data for: BIM-based Intrusion Detection System using Semantic Trajectories for Dynamic Environments**
By **M. Arslan**, C. Cruz, D. Ginhac (2019), Mendeley Data, v1
http://dx.doi.org/10.17632/5hhxtzj5gm.1

Bibliography

(Abramowitz et al., 1972) Abramowitz, M. and Stegun, I.A., Handbook of Mathematical Functions (Dover, New York, 1972).

(Ahn et al., 2017) Ahn, K.U., Kim, D.W., Park, C.S. and De Wilde, P., 2017. Predictability of occupant presence and performance gap in building energy simulation. Applied energy, 208, pp.1639-1652.

(Albanna et al., 2015) Albanna, B.H., Moawad, I.F., Moussa, S.M. and Sakr, M.A., 2015. Semantic trajectories: a survey from modeling to application. In Information Fusion and Geographic Information Systems (IF&GIS'2015) (pp. 59-76). Springer, Cham.

(Ali et al., 2019) Ali, M.U., Hur, S. and Park, Y., 2019. Wi-Fi-based effortless indoor positioning system using IoT sensors. Sensors, 19(7), p.1496.

(Amann et al., 2018) Amann, J., Preidel, C., Tauscher, E. and Borrmann, A., 2018. BIM Programming. In Building Information Modeling (pp. 217-231). Springer, Cham.

(Arslan et al., 2018) Arslan, M., Cruz, C., Roxin, A.M. and Ginhac, D., 2018. Spatio-temporal analysis of trajectories for safer construction sites. Smart and Sustainable Built Environment, 7(1), pp.80-100.

(Arslan et al., 2019a) Arslan, M., Cruz, C. and Ginhac, D., 2019. Semantic enrichment of spatio-temporal trajectories for worker safety on construction sites. Personal and Ubiquitous Computing, pp.1-16. https://doi.org/10.1007/s00779-018-01199-5

(Arslan et al., 2019b) Arslan, M., Cruz, C. and Ginhac, D., 2019. Semantic trajectory insights for worker safety in dynamic environments. Automation in Construction, 106, p.102854. https://doi.org/10.1016/j.autcon.2019.102854

(Arslan et al., 2019c) Arslan, M., Cruz, C. and Ginhac, D., 2019. Visualizing intrusions in dynamic building environments for worker safety. Safety Science, 120, pp.428-446.

(Arslan et al., 2019d) Arslan, M., Cruz, C. and Ginhac, D., 2019, June. DNAS-STriDE Framework for Human Behavior Modeling in Dynamic Environments. In International Conference on Computational Science (pp. 787-793). Springer, Cham. https://doi.org/10.1007/978-3-030-22750-0_79

(Arslan et al., 2019e) Arslan, M., Cruz, C. and Ginhac, D., 2019. Understanding Occupant Behaviors in Dynamic Environments using OBiDE framework. Building and Environment, p.106412. https://doi.org/10.1016/j.buildenv.2019.106412

(Autodesk Dynamo, 2019) The Dynamo primer, 2019, https://primer.dynamobim.org/ (Last accessed 08 Sep. 2019)

(Balzano et al., 2014) Balzano, W. and Del Sorbo, M.R., 2014, September. Setra: A smart framework for gps trajectories' segmentation. In 2014 International Conference on Intelligent Networking and Collaborative Systems (pp. 362-368). IEEE.

(Barthelmes et al., 2017) Barthelmes, V.M., Heo, Y., Fabi, V. and Corgnati, S.P., 2017. Exploration of the Bayesian Network framework for modelling window control behaviour. Building and Environment, 126, pp.318-330.

(Bellamy et al., 2015) Bellamy, L.J., 2015. Exploring the relationship between major hazard, fatal and non-fatal accidents through outcomes and causes. Safety Science, 71, pp.93-103.

(BLS, 2017) Stats.bls.gov.2017, “Census of Fatal Occupational Injuries (CFOI) - current and revised data,” available at: http://stats.bls.gov/iif/oshcfoi1.htm (Last accessed 28 March 2018).

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

(Bozdogan, 1987) Bozdogan, H., 1987. Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. Psychometrika, 52(3), pp.345-370.

(Buchin et al., 2010) Buchin, M., Driemel, A., Van Kreveld, M. and Sacristán, V., 2010, November. An algorithmic framework for segmenting trajectories based on spatio-temporal criteria. In Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems (pp. 202-211). ACM.

(Cai et al., 2018) Cai, G., Lee, K. and Lee, I., 2018. Mining mobility patterns from geotagged photos through semantic trajectory clustering. Cybernetics and Systems, 49(4), pp.234-256.

(Carbonari et al., 2011) Carbonari, A., Giretti, A. and Naticchia, B., 2011. A proactive system for real-time safety management in construction sites. Automation in Construction, 20(6), pp.686-698.

(Carroll, 1995) Carroll, J.M., 1995. Scenario-based design: envisioning work and technology in system development. John Wiley & Sons, Inc..

(Cecconi et al., 2017) Cecconi, F.R., Manfren, M., Tagliabue, L.C., Ciribini, A.L.C. and De Angelis, E., 2017. Probabilistic behavioral modeling in building performance simulation: A Monte Carlo approach. Energy and Buildings, 148, pp.128-141.

(Chaney et al., 2016) Chaney, J., Owens, E.H. and Peacock, A.D., 2016. An evidence based approach to determining residential occupancy and its role in demand response management. Energy and Buildings, 125, pp.254-266.

(Chen et al., 2015) Chen, S., Yang, W., Yoshino, H., Levine, M.D., Newhouse, K. and Hinge, A., 2015. Definition of occupant behavior in residential buildings and its application to behavior analysis in case studies. Energy and Buildings, 104, pp.1-13. https://doi.org/10.1016/j.enbuild.2015.06.075

(Chen et al., 2019) Chen, H., Luo, X., Zheng, Z. and Ke, J., 2019. A proactive workers' safety risk evaluation framework based on position and posture data fusion. Automation in Construction, 98, pp.275-288.

(Cheng et al., 2011) Cheng, T., Venugopal, M., Teizer, J. and Vela, P.A., 2011. Performance evaluation of ultra-wideband technology for construction resource location tracking in harsh environments. Automation in Construction, 20(8), pp.1173-1184.

(Cho et al., 2018) Cho, Y.K., Kim, K., Ma, S. and Ueda, J., 2018. A robotic wearable exoskeleton for construction worker’s safety and health. In ASCE Construction Research Congress (pp. 19-28).

(Choudhry et al., 2008) Choudhry, R.M. and Fang, D., 2008. Why operatives engage in unsafe work behavior: Investigating factors on construction sites. Safety Science, 46(4), pp.566-584.

(Ciabattoni et al., 2019) Ciabattoni, L., Foresi, G., Monteriù, A., Pepa, L., Pagnotta, D.P., Spalazzi, L. and Verdini, F., 2019. Real time indoor localization integrating a model based pedestrian dead reckoning on smartphone and BLE beacons. Journal of Ambient Intelligence and Humanized Computing, 10(1), pp.1-12.

(Ciabattoni et al., 2019) Ciabattoni, L., Foresi, G., Monteriù, A., Pepa, L., Pagnotta, D.P., Spalazzi, L. and Verdini, F., 2019. Real time indoor localization integrating a model based pedestrian dead reckoning on smartphone and BLE beacons. Journal of Ambient Intelligence and Humanized Computing, 10(1), pp.1-12.

(Costin et al., 2015) Costin, A.M., Teizer, J. and Schoner, B., 2015. RFID and BIM-enabled worker location tracking to support real-time building protocol and data visualization. Journal of Information Technology in Construction (ITcon), 20(29), pp.495-517.

(Cruz et al., 2017) Cruz, C., 2017, October. Semantic trajectory modeling for dynamic built environments. In 2017 IEEE International Conference on Data Science and Advanced Analytics (DSAA) (pp. 468-476). IEEE. https://doi.org/10.1109/DSAA.2017.79

(Dabiri et al., 2018) Dabiri, S. and Heaslip, K., 2018. Inferring transportation modes from GPS trajectories using a convolutional neural network. Transportation research part C: emerging technologies, 86, pp.360-371.

(Davis et al., 2010) Davis III, J.A. and Nutter, D.W., 2010. Occupancy diversity factors for common university building types. Energy and buildings, 42(9), pp.1543-1551.

(de Graaff et al., 2016) de Graaff, V., de By, R.A. and van Keulen, M., 2016, April. Automated semantic trajectory annotation with indoor point-of-interest visits in urban areas. In Proceedings of the 31st Annual ACM Symposium on Applied Computing (pp. 552-559). ACM.

(Ding et al., 2013) Ding, L.Y., Zhou, C., Deng, Q.X., Luo, H.B., Ye, X.W., Ni, Y.Q. and Guo, P., 2013. Real-time safety early warning system for cross passage construction in Yangtze Riverbed Metro Tunnel based on the internet of things. Automation in construction, 36, pp.25-37.

(Ding et al., 2018) Ding, L., Fang, W., Luo, H., Love, P.E., Zhong, B. and Ouyang, X., 2018. A deep hybrid learning model to detect unsafe behavior: integrating convolution neural networks and long short-term memory. Automation in Construction, 86, pp.118-124.

(Dong et al., 2017) Dong, X.S., Largay, J.A., Wang, X., Cain, C.T. and Romano, N., 2017. The construction FACE database—Codifying the NIOSH FACE reports. Journal of safety research, 62, pp.217-225.

(Dong et al., 2017) Dong, B. and Biswas, S., 2017. Meal-time and duration monitoring using wearable sensors. Biomedical Signal Processing and Control, 32, pp.97-109.

(Dong et al., 2017) Dong, X.S., Largay, J.A., Wang, X., Cain, C.T. and Romano, N., 2017. The construction FACE database—Codifying the NIOSH FACE reports. Journal of safety research, 62, pp.217-225.

(Du et al., 2018) Du, Z., Yang, Y., Cai, Q., Zhang, C. and Bai, Y., 2018. Modeling and inferring mobile phone users’ negative emotion spreading in social networks. Future Generation Computer Systems, 78, pp.933-942.

(Fang et al., 2016) Fang, Y., Cho, Y.K., Zhang, S. and Perez, E., 2016. Case study of BIM and cloud–enabled real-time RFID indoor localization for construction management applications. Journal of Construction Engineering and Management, 142(7), p.05016003.

(Fang et al., 2018) Fang, W., Ding, L., Luo, H. and Love, P.E., 2018. Falls from heights: A computer vision-based approach for safety harness detection. Automation in Construction, 91, pp.53-61.

(Fang et al., 2018) Fang, Q., Li, H., Luo, X., Ding, L., Rose, T.M., An, W. and Yu, Y., 2018. A deep learning-based method for detecting non-certified work on construction sites. Advanced Engineering Informatics, 35, pp.56-68.

(Feng et al., 2019) Feng, Y. and Golparvar-Fard, M., 2019. Image-Based Localization for Facilitating Construction Field Reporting on Mobile Devices. In Advances in Informatics and Computing in Civil and Construction Engineering (pp. 585-592). Springer, Cham.

(Fileto et al., 2013) Fileto, R., Krüger, M., Pelekis, N., Theodoridis, Y. and Renso, C., 2013, November. Baquara: A holistic ontological framework for movement analysis using linked data. In International Conference on Conceptual Modeling (pp. 342-355). Springer, Berlin, Heidelberg.

(Fileto et al., 2015) Fileto, R., May, C., Renso, C., Pelekis, N., Klein, D. and Theodoridis, Y., 2015. The Baquara2 knowledge-based framework for semantic enrichment and analysis of movement data. Data & Knowledge Engineering, 98, pp.104-122.

(Furletti et al., 2013) Furletti, B., Cintia, P., Renso, C. and Spinsanti, L., 2013, August. Inferring human activities from GPS tracks. In Proceedings of the 2nd ACM SIGKDD International Workshop on Urban Computing (p. 5). ACM.

(Gatto and Jammalamadaka, 2007) Gatto, R. and Jammalamadaka, S.R., 2007. The generalized von Mises distribution. Statistical Methodology, 4(3), pp.341-353.

(Ghosh et al., 2019) Ghosh, A., Chakraborty, A., Chakraborty, D., Saha, M. and Saha, S., 2019. UltraSense: A non-intrusive approach for human activity identification using heterogeneous ultrasonic sensor grid for smart home environment. Journal of Ambient Intelligence and Humanized Computing, pp.1-22.

(Goh et al., 2018) Goh, Y.M., Ubeynarayana, C.U., Wong, K.L.X. and Guo, B.H., 2018. Factors influencing unsafe behaviors: A supervised learning approach. Accident Analysis & Prevention, 118, pp.77-85.

(Golovina et al., 2016) Golovina, O., Teizer, J. and Pradhananga, N., 2016. Heat map generation for predictive safety planning: Preventing struck-by and near miss interactions between workers-on-foot and construction equipment. Automation in construction, 71, pp.99-115.

(Golparvar-Fard et al., 2009) Golparvar-Fard, M., Peña-Mora, F., Arboleda, C.A. and Lee, S., 2009. Visualization of construction progress monitoring with 4D simulation model overlaid on time-lapsed photographs. Journal of computing in civil engineering, 23(6), pp.391-404.

(Gomez-de-Gabriel et al., 2019) Gomez-de-Gabriel, J.M., Fernández-Madrigal, J.A., Lopez-Arquillos, A. and Rubio-Romero, J.C., 2019. Monitoring harness use in construction with BLE beacons. Measurement, 131, pp.329-340.

(Gong et al., 2019) Gong, J., Huang, Y., Chow, P.I., Fua, K., Gerber, M.S., Teachman, B.A. and Barnes, L.E., 2019. Understanding behavioral dynamics of social anxiety among college students through smartphone sensors. Information Fusion, 49, pp.57-68.

(Guo et al., 2017) Guo, H., Yu, Y. and Skitmore, M., 2017. Visualization technology-based construction safety management: A review. Automation in Construction, 73, pp.135-144.

(Guo et al., 2018) Guo, H., Yu, Y., Ding, Q. and Skitmore, M., 2018. Image-and-skeleton-based parameterized approach to real-time identification of construction workers' unsafe behaviors. Journal of Construction Engineering and Management, 144(6), p.04018042.

(Harbelot et al., 2015) Harbelot, B., Arenas, H. and Cruz, C., 2015. 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, pp.3-24.

(Heng et al., 2016) Heng, L., Shuang, D., Skitmore, M., Qinghua, H. and Qin, Y., 2016. Intrusion warning and assessment method for site safety enhancement. Safety Science, 84, pp.97-107.

(Hévízi et al., 2004) Hévízi, G., Biczó, M., Póczos, B., Szabó, Z., Takics, B. and Lorincz, A., 2004, July. Hidden Markov model finds behavioral patterns of users working with a headmouse driven writing tool. In 2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No. 04CH37541) (Vol. 1, pp. 669-674). IEEE.

(Hong et al., 2015) Hong, T., D'Oca, S., Turner, W.J. and Taylor-Lange, S.C., 2015. An ontology to represent energy-related occupant behavior in buildings. Part I: Introduction to the DNAs framework. Building and Environment, 92, pp.764-777.

(Huang and Hinze, 2006) Huang, X. and Hinze, J., 2006. Owner's role in construction safety. Journal of Construction Engineering and Management, 132(2), pp.164-173.

(Huang et al., 2019) Huang, Q., Rodriguez, K., Whetstone, N. and Habel, S., 2019. Rapid Internet of Things (IoT) prototype for accurate people counting towards energy efficient buildings. ITcon, 24, pp.1-13.

(Ilkovičová et al., 2016) Ilkovičová, L., Kajánek, P. and Kopáčik, A., 2016, May. Pedestrian indoor positioning and tracking using smartphone sensors, step detection and map matching algorithm. In International Symposium on Engineering Geodesy (pp. 20-22).

(Jaly-Zada et al., 2015) Jaly-Zada, A., Koch, C. and Tizani, W., 2015. IFC extension for design change management. In Proc., 32nd CIB W78 Conference.

(Jeelani et al., 2018) Jeelani, I., Han, K. and Albert, A., 2018. Automating and scaling personalized safety training using eye-tracking data. Automation in Construction, 93, pp.63-77.

(Jia et al., 2019) Jia, M., Komeily, A., Wang, Y. and Srinivasan, R.S., 2019. Adopting Internet of Things for the development of smart buildings: A review of enabling technologies and applications. Automation in Construction, 101, pp.111-126.

(Ke et al., 2018) Ke, C., Wu, M., Chan, Y. and Lu, K., 2018. Developing a BLE Beacon-Based Location System Using Location Fingerprint Positioning for Smart Home Power Management. Energies, 11(12), p.3464.

(Kontakt, 2019) Kontakt-Beacons, available at https://kontakt.io/ble-beacons-tags/ (2018), Accessed 28th Mar 2018

(Kuhn, 2019) Kuhn, M., 2019. Classification and regression training, https://cran.r-project.org/web/packages/caret/caret.pdf (2019), (Last accessed 27 July 2019)

(Lam et al., 2019) Lam, C.Y. and Ip, W.H., 2019. An Integrated Logistics Routing and Scheduling Network Model with RFID-GPS Data for Supply Chain Management. Wireless Personal Communications, 105(3), pp.803-817.

(Li and Guldenmund, 2018) Li, Y. and Guldenmund, F.W., 2018. Safety management systems: A broad overview of the literature. Safety Science, 103, pp.94-123.

(Li et al., 2015) Li, H., Chan, G., Huang, T., Skitmore, M., Tao, T.Y., Luo, E., Chung, J., Chan, X.S. and Li, Y.F., 2015. Chirp-spread-spectrum-based real time location system for construction safety management: A case study. Automation in Construction, 55, pp.58-65.

(Li et al., 2016) Li, H., Yang, X., Wang, F., Rose, T., Chan, G. and Dong, S., 2016. Stochastic state sequence model to predict construction site safety states through Real-Time Location Systems. Safety science, 84, pp.78-87.

(Li et al., 2018) Li, Y. and Guldenmund, F.W., 2018. Safety management systems: A broad overview of the literature. Safety Science, 103, pp.94-123.

(Liisberg et al., 2016) Liisberg, J., Møller, J.K., Bloem, H., Cipriano, J., Mor, G. and Madsen, H., 2016. Hidden Markov Models for indirect classification of occupant behaviour. Sustainable cities and society, 27, pp.83-98.

(Lim et al., 2016) Lim, J.S., Song, K.I. and Lee, H.L., 2016. Real-time location tracking of multiple construction laborers. Sensors, 16(11), p.1869.

(Liu et al., 2015) Liu, F., Janssens, D., Cui, J., Wets, G. and Cools, M., 2015. Characterizing activity sequences using profile Hidden Markov Models. Expert Systems with Applications, 42(13), pp.5705-5722.

(Liu et al., 2019) Liu, Z., Zhang, C., Guo, Y., Osmani, M. and Demian, P., 2019. A Building Information Modelling (BIM) based Water Efficiency (BWe) Framework for Sustainable Building Design and Construction Management. Electronics, 8(6), p.599.

(Luo et al., 2014) Luo, X., Leite, F. and O'Brien, W.J., 2014. Location-aware sensor data error impact on autonomous crane safety monitoring. Journal of Computing in Civil Engineering, 29(4), p.B4014010.

(Luo et al., 2019) Luo, X., Li, H., Wang, H., Wu, Z., Dai, F. and Cao, D., 2019. Vision-based detection and visualization of dynamic workspaces. Automation in Construction, 104, pp.1-13.

(Mahdavi et al., 2015) Mahdavi, A. and Tahmasebi, F., 2015. Predicting people's presence in buildings: An empirically based model performance analysis. Energy and Buildings, 86, pp.349-355.

(Maltese et al., 2017) Maltese, S., Tagliabue, L.C., Cecconi, F.R., Pasini, D., Manfren, M. and Ciribini, A.L., 2017. Sustainability assessment through green BIM for environmental, social and economic efficiency. Procedia engineering, 180, pp.520-530.

(Markovic et al., 2018) Markovic, R., Grintal, E., Wölki, D., Frisch, J. and van Treeck, C., 2018. Window opening model using deep learning methods. Building and Environment, 145, pp.319-329.

(Marzouk et al., 2018) Marzouk, M. and Al Daour, I., 2018. Planning labor evacuation for construction sites using BIM and agent-based simulation. Safety science, 109, pp.174-185.

(Melfi et al., 2011) Melfi, R., Rosenblum, B., Nordman, B. and Christensen, K., 2011, July. Measuring building occupancy using existing network infrastructure. In 2011 International Green Computing Conference and Workshops (pp. 1-8). IEEE.

(Michelot et al., 2016) Michelot, T., Langrock, R. and Patterson, T.A., 2016. moveHMM: an R package for the statistical modelling of animal movement data using hidden Markov models. Methods in Ecology and Evolution, 7(11), pp.1308-1315.

(Mikkilineni et al., 2019) Mikkilineni, A.K., Dong, J., Kuruganti, T. and Fugate, D., 2019. A novel occupancy detection solution using low-power IR-FPA based wireless occupancy sensor. Energy and Buildings, 192, pp.63-74.

(Mousavi, 2016) Mousavi, A. 2016, "Ontology based semantic knowledge discovery for movement behaviors", doctoral dissertation, University of Calgary.

(Naticchia et al., 2013) Naticchia, B., Vaccarini, M. and Carbonari, A., 2013. A monitoring system for real-time interference control on large construction sites. Automation in Construction, 29, pp.148-160.

(NBS-US, 2019) National BIM Standard (NBS)- United States. Accessed September 30, 2019. https://www.nationalbimstandard.org/faqs.

(Nogueira et al., 2018) Nogueira, T.P., Braga, R.B., de Oliveira, C.T. and Martin, H., 2018. FrameSTEP: A framework for annotating semantic trajectories based on episodes. Expert Systems with Applications, 92, pp.533-545.

(Nogueira, 2017) Nogueira, T.P., 2017. A framework for automatic annotation of semantic trajectories (Doctoral dissertation).

(O'Neill et al., 2013) O'Neill, S., Martinov-Bennie, N., Cheung, A. and Wolfe, K., 2013. Issues in the measurement and reporting of work health and safety performance: a review. Safe Work Australia, Safety Institute of Australia and CPA Australia.

(OSHA, 2019) Occupational safety and health administration (OSHA), 2019, https://www.osha.gov/laws-regs/regulations/standardnumber/1910/1910.145AppA, (Last accessed 08 Sep. 2019)

(Paek et al., 2016) Paek, J., Ko, J. and Shin, H., 2016. A measurement study of BLE iBeacon and geometric adjustment scheme for indoor location-based mobile applications. Mobile Information Systems, 2016.

(Panagiotakis et al., 2011) Panagiotakis, C., Pelekis, N., Kopanakis, I., Ramasso, E. and Theodoridis, Y., 2011. Segmentation and sampling of moving object trajectories based on representativeness. IEEE Transactions on Knowledge and Data Engineering, 24(7), pp.1328-1343.

(Park and Brilakis, 2016) Park, M.W. and Brilakis, I., 2016. Continuous localization of construction workers via integration of detection and tracking. Automation in Construction, 72, pp.129-142.

(Pastell et al., 2018) Pastell, M. and Frondelius, L., 2018. A hidden Markov model to estimate the time dairy cows spend in feeder based on indoor positioning data. Computers and electronics in agriculture, 152, pp.182-185.

(Postawka et al., 2017) Postawka, A., 2017, August. Real-time monitoring system for potentially dangerous activities detection. In 2017 22nd International Conference on Methods and Models in Automation and Robotics (MMAR) (pp. 1005-1008). IEEE.

(Poulos et al., 2018) Poulos, A., Tocornal, F., de la Llera, J.C. and Mitrani-Reiser, J., 2018. Validation of an agent-based building evacuation model with a school drill. Transportation research part C: emerging technologies, 97, pp.82-95.

(Pradhananga and Teizer, 2013) Pradhananga, N. and Teizer, J., 2013. Automatic spatio-temporal analysis of construction site equipment operations using GPS data. Automation in Construction, 29, pp.107-122.

(Rabiner, 1989) Rabiner, L.R., 1989. A tutorial on hidden Markov models and selected applications in speech recognition. Proceedings of the IEEE, 77(2), pp.257-286.

(Rai et al., 2018) Rai, S. and Hu, X., 2018. Building occupancy simulation and data assimilation using a graph-based agent-oriented model. Physica A: Statistical Mechanics and its Applications, 502, pp.270-287.

(Rashid et al., 2017) Rashid, K.M. and Behzadan, A.H., 2017. Risk behavior-based trajectory prediction for construction site safety monitoring. Journal of Construction Engineering and Management, 144(2), p.04017106.

(Riaz et al., 2014) Riaz, Z., Arslan, M., Kiani, A.K. and Azhar, S., 2014. CoSMoS: A BIM and wireless sensor based integrated solution for worker safety in confined spaces. Automation in construction, 45, pp.96-106.

(Ronao et al., 2017) Ronao, C.A. and Cho, S.B., 2017. Recognizing human activities from smartphone sensors using hierarchical continuous hidden Markov models. International Journal of Distributed Sensor Networks, 13(1), p.1550147716683687.

(Rozo et al., 2019) Rozo, K.R., Arellana, J., Santander-Mercado, A. and Jubiz-Diaz, M., 2019. Modelling building emergency evacuation plans considering the dynamic behaviour of pedestrians using agent-based simulation. Safety science, 113, pp.276-284.

(Šabata et al., 2016) Šabata, T., Borovicka, T. and Holena, M., 2016. Modeling and Clustering the Behavior of Animals Using Hidden Markov Models. In CEUR Workshop Proceedings 1649 (pp. 172-178).

(Sankararaman et al., 2013) Sankararaman, S., Agarwal, P.K., Mølhave, T., Pan, J. and Boedihardjo, A.P., 2013, November. Model-driven matching and segmentation of trajectories. In Proceedings of the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (pp. 234-243). ACM.

(Seo, et al., 2015) Seo, J., Han, S., Lee, S. and Kim, H., 2015. Computer vision techniques for construction safety and health monitoring. Advanced Engineering Informatics, 29(2), pp.239-251.

(Shtar et al., 2019) Shtar, G., Shapira, B. and Rokach, L., 2019. Clustering Wi-Fi fingerprints for indoor–outdoor detection. Wireless Networks, 25(3), pp.1341-1359.

(Shuang et al., 2019) Shuang, D., Heng, L., Skitmore, M. and Qin, Y., 2019. An experimental study of intrusion behaviors on construction sites: The role of age and gender. Safety science, 115, pp.425-434.

(Sokolova and Lapalme, 2009) Sokolova, M. and Lapalme, G., 2009. A systematic analysis of performance measures for classification tasks. Information processing & management, 45(4), pp.427-437.

(Son et al., 2019) Son, H., Choi, H., Seong, H. and Kim, C., 2019. Detection of construction workers under varying poses and changing background in image sequences via very deep residual networks. Automation in Construction, 99, pp.27-38.

(Spaccapietra et al., 2008) Spaccapietra, S., Parent, C., Damiani, M.L., de Macedo, J.A., Porto, F. and Vangenot, C., 2008. A conceptual view on trajectories. Data & knowledge engineering, 65(1), pp.126-146.

(Stojanovic et al., 2018) Stojanovic, V., Trapp, M., Richter, R., Hagedorn, B. and Döllner, J., 2018. Towards The Generation of Digital Twins for Facility Management Based on 3D Point Clouds. In Proceedings of the 34th Annual ARCOM Conference.

(Tang et al., 2019) Tang, S., Shelden, D.R., Eastman, C.M., Pishdad-Bozorgi, P. and Gao, X., 2019. A review of building information modeling (BIM) and the internet of things (IoT) devices integration: Present status and future trends. Automation in Construction, 101, pp.127-139.

(Teizer and Cheng, 2015) Teizer, J. and Cheng, T., 2015. Proximity hazard indicator for workers-on-foot near miss interactions with construction equipment and geo-referenced hazard areas. Automation in Construction, 60, pp.58-73.

(Teizer and Cheng, 2015) Teizer, J. and Cheng, T., 2015. Proximity hazard indicator for workers-on-foot near miss interactions with construction equipment and geo-referenced hazard areas. Automation in Construction, 60, pp.58-73.

(Teizer et al., 2013) Teizer, J., Cheng, T. and Fang, Y., 2013. Location tracking and data visualization technology to advance construction ironworkers' education and training in safety and productivity. Automation in Construction, 35, pp.53-68.

(Teizer et al., 2015) Teizer, J. and Cheng, T., 2015. Proximity hazard indicator for workers-on-foot near miss interactions with construction equipment and geo-referenced hazard areas. Automation in Construction, 60, pp.58-73.

(Ulmeanu et al., 2017) Ulmeanu, A.P., Barbu, V.S., Tanasiev, V. and Badea, A., 2017. Hidden Markov Models revealing the household thermal profiling from smart meter data. Energy and Buildings, 154, pp.127-140.

(Van der Heijden, 2011) Van der Heijden, K., 2011. Scenarios: the art of strategic conversation. John Wiley & Sons.

(Volk et al., 2014) Volk, R., Stengel, J. and Schultmann, F., 2014. Building Information Modeling (BIM) for existing buildings—Literature review and future needs. Automation in Construction, 38, pp.109-127.

(Wagner et al., 2018) Wagner, A., O'Brien, W. and Dong, B., 2018. Exploring Occupant Behavior in Buildings. Wagner, A., O'Brien, W., Dong, B., Eds. https://doi.org/10.1007/978-3-319-61464-9

(Wan et al., 2018) Wan, C., Zhu, Y., Yu, J. and Shen, Y., 2018. SMOPAT: Mining semantic mobility patterns from trajectories of private vehicles. Information Sciences, 429, pp.12-25.

(Wang et al., 2017) Wang, Z.Z. and Chen, C., 2017. Fuzzy comprehensive Bayesian network-based safety risk assessment for metro construction projects. Tunnelling and Underground Space Technology, 70, pp.330-342.

(Wang et al., 2018) Wang, T., Qiao, M., Deng, Y., Zhou, Y., Wang, H., Lyu, Q. and Snoussi, H., 2018. Abnormal event detection based on analysis of movement information of video sequence. Optik, 152, pp.50-60.

(Wang et al., 2019) Wang, D., Li, H. and Chen, J., 2019. Detecting and measuring construction workers' vigilance through hybrid kinematic-EEG signals. Automation in Construction, 100, pp.11-23.

(Williams et al., 2017) Williams, M.L., James, W.P. and Rose, M.T., 2017. Fixed-time data segmentation and behavior classification of pasture-based cattle: Enhancing performance using a hidden Markov model. Computers and electronics in agriculture, 142, pp.585-596.

(Wolf et al., 2019) Wolf, S., Møller, J.K., Bitsch, M.A., Krogstie, J. and Madsen, H., 2019. A Markov-Switching model for building occupant activity estimation. Energy and Buildings, 183, pp.672-683.

(Wu et al., 2010) Wu, W., Yang, H., Chew, D.A., Yang, S.H., Gibb, A.G. and Li, Q., 2010. Towards an autonomous real-time tracking system of near-miss accidents on construction sites. Automation in Construction, 19(2), pp.134-141.

(Wu et al., 2010) Wu, W., Yang, H., Chew, D.A., Yang, S.H., Gibb, A.G. and Li, Q., 2010. Towards an autonomous real-time tracking system of near-miss accidents on construction sites. Automation in Construction, 19(2), pp.134-141.

(Wu et al., 2013) Wu, W., Yang, H., Li, Q. and Chew, D., 2013. An integrated information management model for proactive prevention of struck-by-falling-object accidents on construction sites. Automation in Construction, 34, pp.67-74.

(Wu et al., 2015) Wu, F., Li, Z., Lee, W.C., Wang, H. and Huang, Z., 2015, May. Semantic annotation of mobility data using social media. In Proceedings of the 24th International Conference on World Wide Web (pp. 1253-1263). International World Wide Web Conferences Steering Committee.

(Wu et al., 2015) Wu, F., Wang, H., Li, Z., Lee, W.C. and Huang, Z., 2015, May. SemMobi: A semantic annotation system for mobility data. In Proceedings of the 24th International Conference on World Wide Web (pp. 255-258). ACM.

(Yan et al., 2011) Yan, Z., Chakraborty, D., Parent, C., Spaccapietra, S. and 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.

(Yan et al., 2011) Yan, Z., Giatrakos, N., Katsikaros, V., Pelekis, N. and Theodoridis, Y., 2011, August. SeTraStream: semantic-aware trajectory construction over streaming movement data. In International Symposium on Spatial and Temporal Databases (pp. 367-385). Springer, Berlin, Heidelberg.

(Yan et al., 2013) 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.

(Yan et al., 2015) Yan, D., O'Brien, W., Hong, T., Feng, X., Gunay, H.B., Tahmasebi, F. and Mahdavi, A., 2015. Occupant behavior modeling for building performance simulation: Current state and future challenges. Energy and Buildings, 107, pp.264-278. https://doi.org/10.1016/j.enbuild.2015.08.032

(Yang et al., 2012) Yang, H., Chew, D.A., Wu, W., Zhou, Z. and Li, Q., 2012. Design and implementation of an identification system in construction site safety for proactive accident prevention. Accident Analysis & Prevention, 48, pp.193-203.

(Yang et al., 2017) Yang, X., Luo, X., Li, H., Luo, X. and Guo, H., 2017. Location-based measurement and visualization for interdependence network on construction sites. Advanced Engineering Informatics, 34, pp.36-45.

(Yang et al., 2019) Yang, K. and Ahn, C.R., 2019. Inferring workplace safety hazards from the spatial patterns of workers' wearable data. Advanced Engineering Informatics, 41, p.100924.

(Z. Yan, 2011) Yan, Z., 2011. Semantic trajectories: computing and understanding mobility data (Doctoral dissertation, Verlag nicht ermittelbar).

(Žalik et al., 2008) Žalik, K.R., 2008. An efficient k′-means clustering algorithm. Pattern Recognition Letters, 29(9), pp.1385-1391.

(Zhang et al., 2015) Zhang, S., Teizer, J., Pradhananga, N. and Eastman, C.M., 2015. Workforce location tracking to model, visualize and analyze workspace requirements in building information models for construction safety planning. Automation in Construction, 60, pp.74-86.

(Zhang et al., 2018) Zhang, M., Cao, T. and Zhao, X., 2018. Using Smartphones to Detect and Identify Construction Workers' Near-Miss Falls Based on ANN. Journal of Construction Engineering and Management, 145(1), p.04018120.

(Zhang et al., 2019) Zhang, H., Yan, X., Li, H., Jin, R. and Fu, H., 2019. Real-Time Alarming, Monitoring, and Locating for Non-Hard-Hat Use in Construction. Journal of Construction Engineering and Management, 145(3), p.04019006.

(Zhang et al., 2019) Zhang, P., Li, N., Jiang, Z., Fang, D. and Anumba, C.J., 2019. An agent-based modeling approach for understanding the effect of worker-management interactions on construction workers' safety-related behaviors. Automation in Construction, 97, pp.29-43.

(Zheng, 2015) Zheng, Y., 2015. Trajectory data mining: an overview. ACM Transactions on Intelligent Systems and Technology (TIST), 6(3), p.29.

(Zhou et al., 2019) Zhou, Z., Li, C., Mi, C. and Qian, L., 2019. Exploring the potential use of near-miss information to improve construction safety performance. Sustainability, 11(5), p.1264.

(Zhu et al., 2016) Zhu, Z., Park, M.W., Koch, C., Soltani, M., Hammad, A. and Davari, K., 2016. Predicting movements of onsite workers and mobile equipment for enhancing construction site safety. Automation in Construction, 68, pp.95-101.

(Zin et al., 2016) Zin, T.T., Kobayashi, I., Tin, P. and Hama, H., 2016, May. A general video surveillance framework for animal behavior analysis. In 2016 Third International Conference on Computing Measurement Control and Sensor Network (CMCSN) (pp. 130-133). IEEE.

(Zucchini, 2000) Zucchini, W., 2000. An introduction to model selection. Journal of mathematical psychology, 44(1), pp.41-61.