DTU Library

A Data-Driven Method for Hazard Zone Identification in Construction Sites with Wearable Sensors

Hong, Kepeng; Teizer, Jochen

Published in:
Proceedings of the CIBW099W123

Publication date:
2023

Document Version
Publisher's PDF, also known as Version of record

Citation (APA):
Hong, K., & Teizer, J. (2023). A Data-Driven Method for Hazard Zone Identification in Construction Sites with Wearable Sensors. In E. Fidelis, F. Sherratt, & A. Soeiro (Eds.), *Proceedings of the CIBW099W123: Digital Transformation of Health and Safety in Construction* (pp. 41-48) https://doi.org/10.24840/978-972-752-309-2

A DATA-DRIVEN METHOD FOR HAZARD ZONE IDENTIFICATION IN CONSTRUCTION SITES WITH WEARABLE SENSORS

Kepeng Hong, Jochen Teizer

Dept. of Civil and Mechanical Engineering, Technical University of Denmark (DENMARK)

Abstract

Hazard zone identification plays a significant role in designing construction site layouts and preventing construction accidents. The existing identification methods require laborious and empirical predefinitions. Despite the emergence of some automatic hazard zone identification algorithms, they still heavily depend on stagnant regulations and rules. Thanks to the development of real-time location systems and sensor technology, the trajectory data of construction workers and equipment can be precisely collected and stored. Such data can facilitate understanding workers' and equipment's activity patterns, thus further improving the dynamic recognition of hazardous behaviors and areas. In this work, we introduce a data-driven method to automatically identify and predict potential hazard zones in the construction site. The algorithm is implemented on a digital twin platform to retrieve location data and generate real-time hazard index maps. The method consists of the following parts: (a) construction site sensor data collection and processing, (b) worker and equipment data analysis (e.g., speed, acceleration, and trajectory), and (c) hazard zone identification algorithms development. For validation, we implement the method on one railway construction project in Karlsruhe and compare the result with the close-call incidents map. This real-life case study partially demonstrates the effectiveness and accuracy of our method under the constraints of currently limited project data. On the basis of this work, further study can be conducted on the aspects of workers' behavioral patterns and prediction model selection.

Keywords: construction safety, data-driven method, digital twins, hazard zone identification, real-time location sensing.

1 INTRODUCTION

Construction sites are widely recognized as one of the most dangerous workplaces. The rates of fatality and non-fatality accidents are significantly higher than in other industries. A report on the labor statistics in the EU shows that more than a fifth of all fatal accidents took place in the EU within the construction sector of which, depending the country, up to 15 percent relate struck-by equipment or other object incidents [1]. Traditionally, Health, Safety, and Environmental (HSE) managers define the construction hazard zone in advance and observe the workers' safety performance based on the predefined hazard zone. However, the construction site's dynamic and changeable features bring challenges to the static and top-down approach to delineating the hazard zone area.

The emergence of the Digital Twin (DT) even in construction allows – at some point in future time – HSE managers to proactively monitor and interfere with potential hazards at construction sites in real-time [2]. A DT can first be a virtual representation of a real-world product or system (aka. physical twin). By integrating data from the Internet of Things (IoT) with Building Information Modeling (BIM), project stakeholders can add valuable information to a Digital Twin Platform (DTP) of the construction site for viewing and controlling the on-site events and processes. Through various sensing technologies, such as Real-time Location Sensing (RTLS) [3, 4] or vision-based sensors [5], the DT can retrieve contextual information. These are, for example, trajectories of resources, incl. workers or equipment, or image-based point clouds of a work terrain they traverse in. Such (trajectory) data can then be analyzed using hazard detection algorithms that predict future work states, and in case applicable, avoid potentially dangerous situations. These algorithms pose a research gap, because they rely so far on basic data analysis and visualization [5]. Instead, proposed are additional criteria, such as the capitalization of a priori-known safety rules [6] or real-time data of too close proximity distances [7]. Such algorithms can then be powerful when implemented on a DTP as they can provide real-time personalized feedback to the HSE managers, equipment operators, and workers, and proactively detect and avoid personal harm or other collateral damage.

Nevertheless, a DTP still heavily relies on predefined input (e.g., safety distance, hazard zone predefinition) from HSE managers, which undermines the effectiveness of real-time hazard detection and feedback. At the same time, spatial-temporal data from workers can reveal their behavioral patterns related to surrounding contexts. Therefore, we aim to utilize the historical and real-time workers' location data on the DTP for Hazard Zone Identification (HZI) so that the predefined hazard zone can be adjusted on time. Previous studies have utilized the location data or behavioral patterns of construction workers to analyze the safety condition at the construction site. For instance, Li et al. apply a crowd-sourced density map method to classify construction site zones for workplace safety [8]. Yang et al. analyze abnormal patterns of workers' gait cycles to infer workplace safety hazards [9]. Golovina et al. use heatmap for recording, identifying, and analyzing interactive hazardous near-miss situations between workers on foot and heavy construction equipment [10]. So far, these approaches only focus on one certain aspect of the temporal-spatial data for hazard prediction and have only been tested on locally stored data, without further integration with real-time data on a DTP.

In this work, we focus on developing a method of updating hazard zone maps using only location data from construction workers. We derive the visiting frequency, speed patterns, and stay duration for each spot from the aggregate spatiotemporal data of workers. With the fusion of the visiting frequency, speed patterns, and stay duration, we calculate the hazard index for each spot. The algorithm is expected to be integrated into DTP so that it can fetch real-time location data and predict the hazard index map. In this work, we use the persistence model for hazard prediction for simplicity.

Our contributions are listed as follows:

- We establish a framework of a DTP for construction site hazard zone prediction using only location data from the construction workers. The whole process of data collection, communication, analysis, and feedback is defined in the framework.
- We analyze location data (e.g., visiting frequency, speed abnormality, and stay duration), revealing human behavioral patterns to infer the contextual workplace safety conditions.
- We develop an algorithm for hazard zone prediction using a persistence model. The algorithm uses the spatial-temporal information of the construction workers as input data. It derives a hazard index for each spot, which can be used to supplement a predefined hazard zone map.

2 METHODS

In this work, we conduct a case study on a railway construction project at Karlsruhe. As shown in Figure 1, the method adopted in this study consists of three consecutive steps. First, we define the framework and workflow of DTP. Subsequently, we analyze and conclude the behavioral patterns of workers from the location and vision-based data, based on which we develop the algorithm for the HZI. Lastly, we implement the algorithm using the dataset to generate the hazard index map. The hazard index map is further validated via comparison with the close call incidents occurrence map. Close calls refer to incidents that nearly turned into an accident, which is defined by Occupational Safety and Health Administration in 2016 [11]. Proximity analysis is a specific method to detect close call events at the construction site [7]. When workers are too close to the path of the heavy equipment, the proximity-based events will be recorded in terms of location and timestamp. Despite various types of hazards (e.g., falls for lack of guard rails, struck-by due to obstacles), we can only partially validate the results of the hazard zone using the proximity-based close call analysis with the mere input of location data.

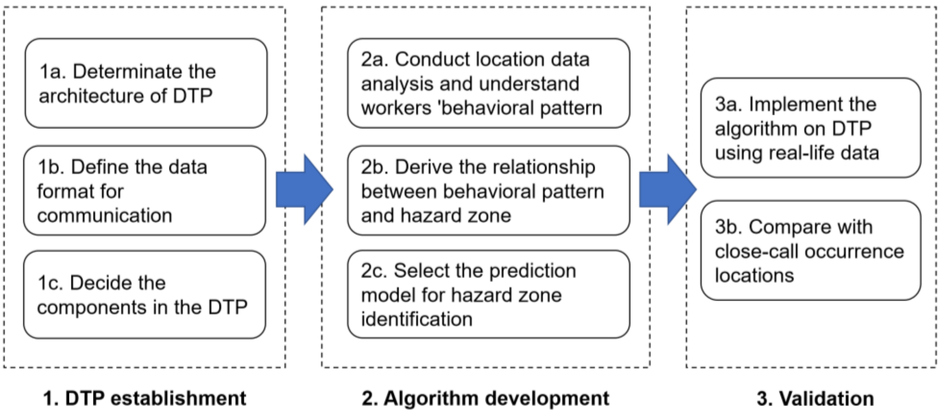

Figure 1: Methodology for the development of DTP and HZI algorithm.

3 DIGITAL TWIN PLATFORM

DTP is where the on-site location data is received and processed. We define the overall framework, data format, and components of the DTP.

3.1 Framework

Figure 2 presents the overall framework and workflow of receiving data, processing data, and returning feedback on the DTP. The framework comprises three parts: real-time location data collection at the construction site, data preprocessing, and real-time hazard map generation. The RTLS system at the construction site reports all relevant event data (e.g., equipment ID, personal tag ID, location coordinate, timestamp) to the DTP. After the DTP receives the data, preprocessing module (PPM) reduces the noise in the data, converts the geographic coordinate to the local coordinate, and derives other required information (e.g., speed). The preprocessed data is the input for the HZI algorithm, which is elaborated on in the next section.

We adopt the Message Queuing Telemetry Transport (MQTT) protocol as the data communication protocol in the DTP. MQTT is an OASIS standard messaging protocol for the Internet of Things. It is designed as an extremely lightweight publish/subscribe messaging transport that is ideal for connecting remote devices with a small code footprint and minimal network bandwidth [12].

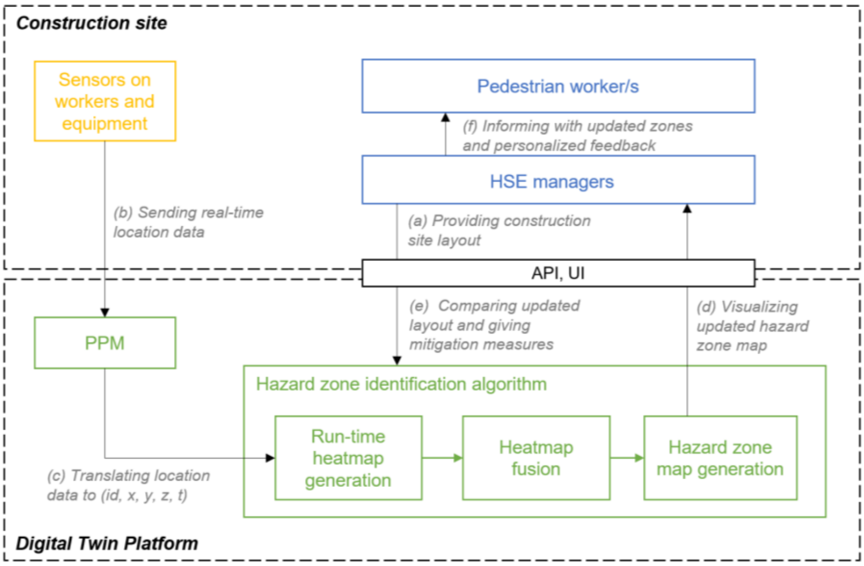

Figure 2: Framework and workflow of receiving and processing location data on the DTP.

3.2 Components

The IoT component of the DTP captures data from the physical objects in the real world. Continuously recording data from sensor tags, such as Real-time Kinematic Global Navigation Satellite System (RTK-GNSS), placed or installed on pedestrian workers and dangerous equipment, can yield a valuable source of spatial and temporal information that is needed for analyzing too close proximity events of such objects to each other.

For instance, the Karlsruhe railway construction project applied RTK-GNSS sensor system consisting of a base station and rovers for real-time resource location data collection. The employed RTK-GNSS can provide cm-level location accuracy, which is essential for monitoring workers' safety at construction sites [13].

As a further element of the DTP, the information displayed on a User Interface (UI) should be designed distinctively for multi-level users. These include but are not limited to equipment operators, pedestrian workers, and HSE managers. HSE managers, for example, may have access to the real-time trajectory data for each piece of equipment and workers and a full analysis of the recorded raw data. In contrast, workers should have access to only their trajectory and their personalized feedback.

The DTP therefore consists of modules responsible for data gathering, processing, analysis, and visualization. The functionalities of the different modules are listed as follows:

- a data preprocessing module that translates data from the RTLS system into the input data for the HZI algorithms,
- a proactive safety monitoring module that analyses trajectory and heatmap in the work environment, where the HZI algorithm is implemented, and
- a UI module that visualizes the safety analysis provides personalized feedback and assists decision-making processes.

4 HAZARD ZONE IDENTIFICATION ALGORITHM

The heatmap is often used to compare with the predefined construction site hazard map and analyze the visiting frequency of different spots in the construction site. Figure 3 shows an example of visiting frequency heatmap, where the construction site is divided into 1 m x 1 m cells, and the count of unique visits is shown inside each cell. The visiting frequency can indicate the hazard degree of each spot at the construction site, i.e., the cells that have been visited more tend to have less hazard degree.

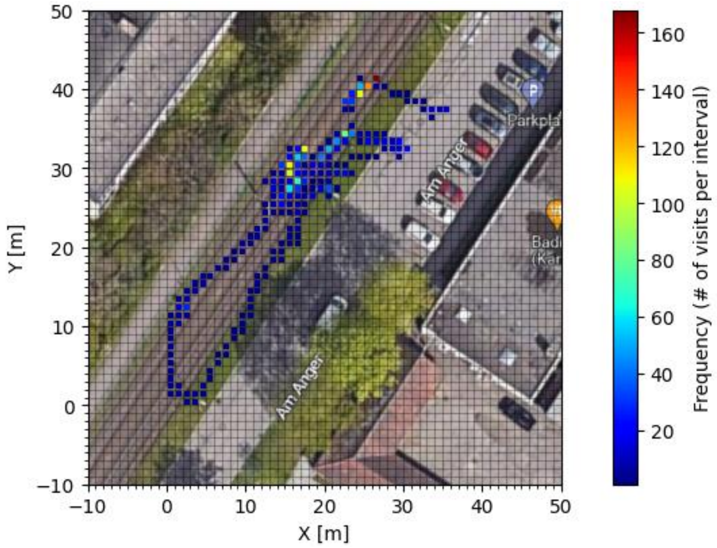

Figure 3: Frequency heatmap of all workers visiting the spots at the construction site (the site is divided into an occupancy grid of 1 m x 1 m cells).

The HZI algorithm uses real-time location data as input and generates the hazard zone map after each time interval. We initially observe and analyze the speed patterns of the workers at the construction site. Figure 4 shows that the speed of construction workers is within a range between 0 and 0.68 m/s. However, we notice that there are two noticeable speed changes during the observation. In this work, we define a speed above 0.68 m/s as speed abnormality. We investigated when and where the changes occurred based on the frames captured by the surveillance camera on the construction sites. As shown in Figure 5, at T = 50 s and T = 70 s, the workers change their speed to keep their distance from the path of the concrete truck. These speed abnormalities tend to occur when workers enter a hazardous area or when the equipment is close to workers. Therefore, it is plausible to infer a correlation between the worker's speed abnormality and their proximity to the hazard area.

In addition to the speed abnormality, maximum stay duration at one region can also indicate the safety condition of the region. From the observation and inference, we conclude the following hypotheses,

Hypothesis 1: Construction workers, after safety training, can recognize and avoid the hazards at the construction site. Given the dynamic feature of construction sites, the real-time location information of construction workers can convey the latest information regarding the hazard zones at the construction site, which still needs to be identified at the predefined construction hazard zones.

Hypothesis 2: Construction workers conducting the same types of work tend to follow the same routes and behavioral patterns. When they get close to hazard zones, they respond similarly, such as adjusting speed and reducing stay duration.

Hypothesis 3: We use the persistence model for hazard zone prediction in the algorithm, which assumes the consistent state of the construction site for the adjunct time intervals.

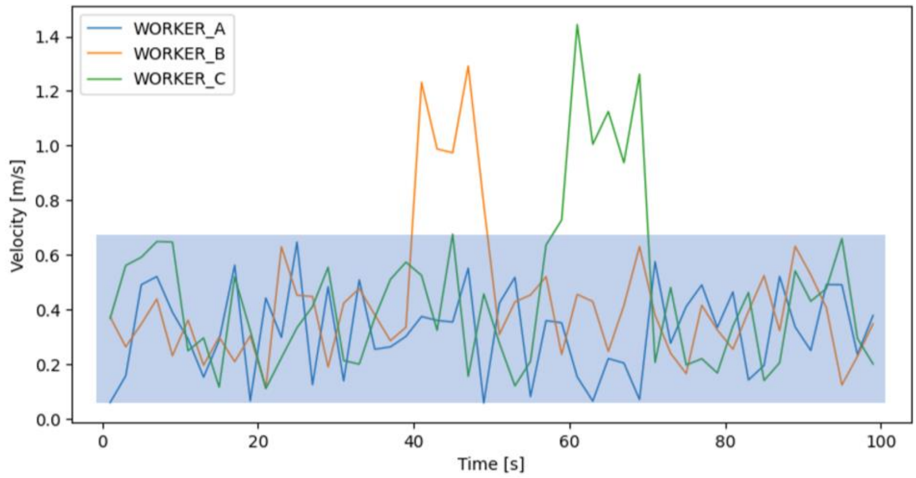

Figure 4: Automatically generated speed patterns of the construction workers.

Figure 5: Visualization of the ‘speed abnormality’ (note: manual annotation of colors).

Figure 6 explains the workflow of the HZI algorithm. The construction site is divided into cells, and the algorithm aggregates the spatial and temporal data from the workers in the cell. For each cell (C_{ij}), the visiting frequency (VF), speed abnormality count (SA), and maximum stay duration (SD) are added up and normalized over all the cells at the construction site. For each cell visited, we calculate the hazard index (HI) of the cell (C_{ij}) at a given time T. The hazard index is then normalized over all cells. The higher hazard index indicates a higher incident risk in the area.

$$HI_{C_{ij}}^{T} = \frac{\sum_{T-\Delta T}^{T}(SA_n - SD_n)}{VF_n}$$

With the persistence model, the hazard index for each cell at the next interval ($T, T + \Delta T$) is considered the same as the hazard index at the given time T.

$$HI(T, T + \Delta T) = HI(T)$$

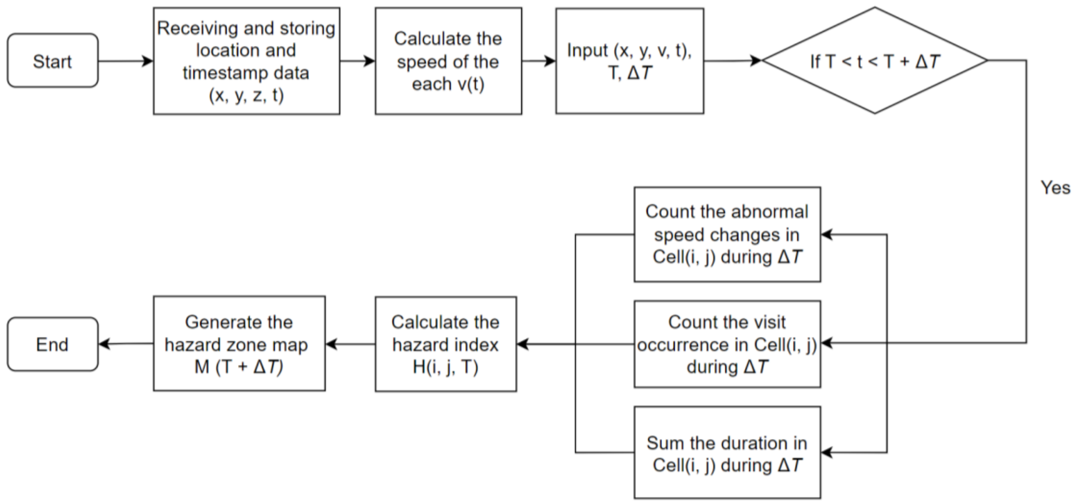

Figure 6: Hazard zone identification algorithm workflow.

5 RESULTS

The dataset for the Karlsruhe project was collected by Real-time Kinematic-Global Navigation Satellite System (RTK-GNSS) modules, which contain location and timestamp data of four tags, respectively, one concrete truck and three workers. As shown in Table 1, the entire dataset spans fifteen minutes, with each tag sending its latitude, longitude, and elevation information every second. A close call analysis was conducted in the previous study, and the close call occurrence locations can be compared with the algorithm's hazard zone map.

Table 1: Overview of an extracted part of the dataset collected from the Karlsruhe railway project site.

Identification	Count (No.)	Start time	End time	Date
WORKER_A	910	13:05:09	13:20:18	09-Jul-2022
WORKER_B	910	13:05:09	13:20:18	09-Jul-2022
WORKER_C	910	13:05:09	13:20:18	09-Jul-2022
CONCRETE_TRUCK_A	910	13:05:09	13:20:18	09-Jul-2022

We simulate the real-time data communication process using an MQTT broker and establish a DTP on a local server. We implemented the HZI algorithm on the data collected at the Karlsruhe railway construction project. The location and timestamp data of three tags are used as input, and the time interval ΔT is set as 10 minutes. Figure 7 shows the input data, the intermediate heatmaps, and the hazard index map generated from the algorithm.

The close call incidents are based on the proximity between workers and heavy machinery, which can be caused by unscrutinized layout arrangement of the construction site or latent surrounding hazards (e.g., obstacles on the pedestrian's pathway). Therefore, a close call incidence map can indicate the potential hazard spots at the construction site. We adopt the close call detection and analysis algorithm developed by Golovina et al. [7]. When the workers are within the protective envelope of the equipment, the events will be recorded. In our case study, we define a proximity-based close call event as when workers are within the one-meter protective envelope outside the vehicle, as shown in Fig. 5.

We compare the hazard index map with the close call incidence map during the interval. As shown in Figure 8, the comparison indicates that the location of close call incidents is around the high hazard index area. However, the locations of close call incidents only partially overlap with the area with a higher hazard index, which can result from factors such as the delay of worker reaction. Due to the limited amount of data, bias exists in both the hazard index map and the close call incident map. Further algorithm validation should be carried out on the dataset with broader time windows and diverse data types.

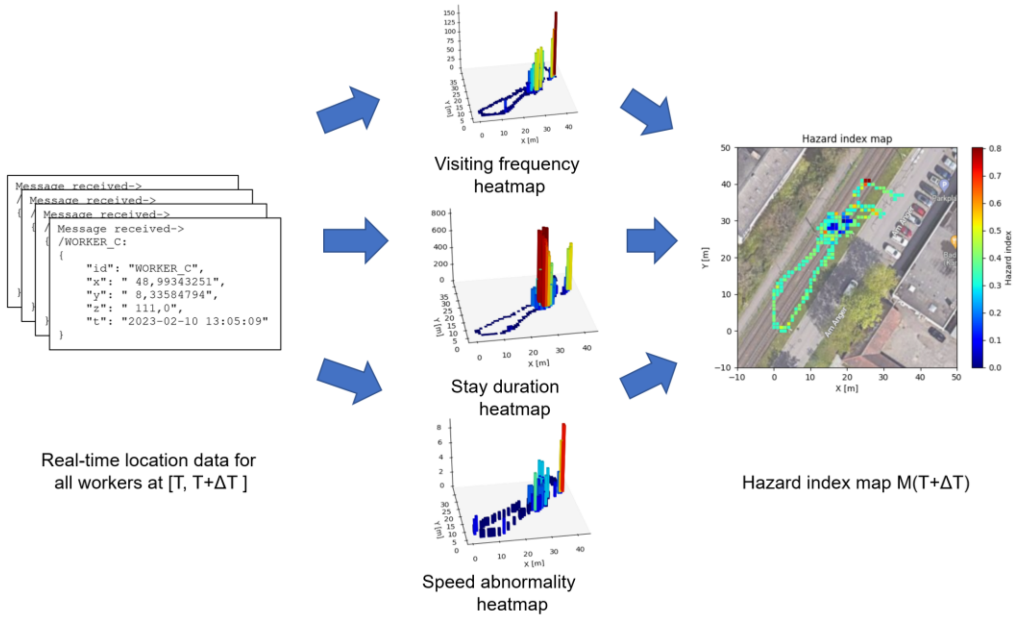

Figure 7: Results from the hazard zone identification algorithm.

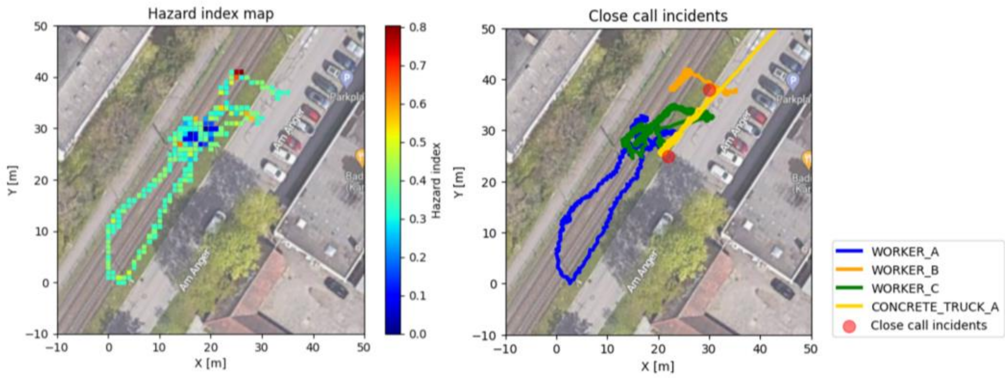

Figure 8: Comparison between the hazard index map and close call incident map.

6 DISCUSSION

The novelty of this work is to propose an algorithm to identify and predict the hazard zone at the construction site based on behavioral patterns derived from workers' spatial-temporal data. Implemented on the DTP, the algorithm can process on-site data and return results to the HSE managers in real time to support hazard meditation and decision-making. Due to the constraint of limited data availability, the accuracy and effectiveness of the algorithm can only be partially validated in this study. The HZI algorithm must be further validated with expert knowledge and other hazard detection methods, such as BIM- and vision-based methods.

The algorithm is built on the hypotheses as mentioned above and assumptions. Improvements in the algorithm are expected in the aspects of the prediction model and semantic enrichment. The prototypical algorithm uses using persistence model for simplicity, which assumes the consistent state of the construction site in the adjunct time intervals. The prediction model should include historical data for comprehensive assessment and robust performance in further research. It requires access to historical data storage on the DTP.

Regarding semantic enrichment, additional data can be integrated into the algorithm. Supplementary data, such as work type and experienced level, can help us better classify the behavioral patterns of construction workers. For instance, workers of certain work types should follow similar routes and schedules, whose trajectories can be grouped for investigation. Basic contextual information can also semantically facilitate our understanding of trajectories so that we can categorize the zones at the construction site. Workers' behavior is expected to differ in the rest area and working area, and the algorithm is supposed to focus on the working area.

7 CONCLUSION

In this work, we developed the concept of a DTP and implemented RTK-GNSS an automated HZI function using real-time location data from the construction site. Using spatial-temporal information, the algorithm can derive construction workers' general behavioral patterns (e.g., speed and stay duration), from which we can infer the hazard status of the surrounding environment. The prototypical algorithm fuses visiting frequency, stay duration, and speed abnormality for the prediction of the hazard index at each spot on the construction site using a persistence model.

We test the DTP and algorithm using the data from a real-life railway construction project. The framework and workflow prove to be viable. We compare the hazard index map with the location of close call incidents that occurred during the time interval. However, under the constraint of limited data availability, the hazard index map generated from the algorithm can only be partially validated compared with a proximity-based close call occurrence map. Further validation is expected with more input data and other hazard identification approaches.

8 ACKNOWLEDGEMENT

The research presented in this paper has been funded by the European Union Horizon 2020 research and innovation program under grant agreement no. 958310.

REFERENCES

[1] Eurostat. (2022, October). Accidents at work - statistics by economic activity. From https://ec.europa.eu/eurostat/statistics-explained/index.php.

[2] Teizer, J., Johansen, K. W., Schultz, C. (2022). “The Concept of Digital Twin for Construction Safety”. Construction Research Congress 2022, 1156–1165. https://doi.org/10.1061/9780784483961.121

[3] Oloufa, A., Ikeda, M., Oda, H. (2003). “GPS-Based wireless collision detection of construction equipment”. NIST Special Publication, 461-466. https://doi.org/10.22260/ISARC2002/0072.

[4] Cheng, T., Migliaccio, G.C., Teizer J., Gatti, U.C. (2013). “Data Fusion of Real-time Location Sensing (RTLS) and Physiological Status Monitoring (PSM) for Ergonomics Analysis of Construction Workers”, Journal of Computing in Civil Engineering, ASCE, 27(3), 320-335, http://dx.doi.org/10.1061/(ASCE)CP.1943-5487.0000222.

[5] Teizer, J. (2015). “Status quo and open challenges in vision-based sensing and tracking of temporary resources on infrastructure construction sites”. Advanced Engineering Informatics, 29(2), 225-238. https://doi.org/10.1016/j.aei.2015.03.006.

[6] Schwabe, K., Teizer, J., König, M. (2019). “Applying rule-based model-checking to construction site layout planning tasks”, Automation in Construction, Elsevier, 97, 205-219, https://doi.org/10.1016/j.autcon.2018.10.012.

[7] Golovina, O., Perschewski, M., Teizer, J., König, M. (2019). “Algorithm for Quantitative Analysis of Close Call Events and Personalized Feedback in Construction Safety”, Automation in Construction, 99, 206-222, https://doi.org/10.1016/j.autcon.2018.11.014.

[8] Li, H., Yang, X., Skitmore, M., Wang, F., Forsythe, P. (2017). “Automated classification of construction site hazard zones by crowd-sourced integrated density maps”. Automation in Construction, Elsevier, 81, 328–339. https://doi.org/10.1016/j.autcon.2017.04.007.

[9] Yang, K., Ahn, C. R. (2019). “Inferring workplace safety hazards from the spatial patterns of workers’ wearable data. Advanced Engineering Informatics”, 41, 100924.

[10] Golovina, O., Teizer, J., 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, Elsevier, 71, 99-115, http://dx.doi.org/10.1016/j.autcon.2016.03.008.

[11] Occupational Safety and Health Administration. (2016). “Recommended Practices for Safety and Health Programs in Construction”, from https://www.osha.gov/sites/default/files/OSHA3886.pdf

[12] Katsikeas, S. et al., "Lightweight & secure industrial IoT communications via the MQ telemetry transport protocol," 2017 IEEE Symp. on Computers and Communications, 2017, pp. 1193-1200.

[13] Varbla, S., Puust, R., Ellmann R. (2021) “Accuracy assessment of RTK-GNSS equipped UAV conducted as-built surveys for construction site modelling”, Survey Review, 53:381, 477-492.