Towards data mining on construction sites: Heterogeneous data acquisition and fusion

F. Pfitzner, A. Braun & A. Borrmann
Chair of Computational Modeling and Simulation
Technical University of Munich, Munich, Germany

ABSTRACT: Data mining methods can invoke substantial optimization potential, as demonstrated in the manufacturing industry. Looking at the construction industry and precisely the as-performed stage, though, this research area is in its infancy worldwide. By now, it is not clear how specific on-site activities can be monitored adequately and how diverse data sources can be combined to make transparent and invulnerable statements about particular on-site activities. The presented study investigates the application of modern data analysis methods to ongoing construction projects to reveal information about specific activities. Raw data from various construction sites has been gained using camera systems, Bluetooth Low Energy (BLE) sensors, and laser scanners to build a powerful foundation of data sources. A data pipeline for integrating different data sources has been developed to handle a large amount and variety of data and its subsequent processing into higher-level information. Using a data mining approach, namely map-reduce, we scaled the significant amount of data down to particular problem-targeted databases. Object detection methods were applied to process images of the construction activities. It was possible to detect on-site construction workers' start and end times, breaks, and location. The introduced results have been verified by using fixed beacons and heterogeneous data types. In conclusion, the presented research provides fundamental methods for examining existing construction processes and collecting data for future analyses.

1 INTRODUCTION

Data mining (DM) has been used in many industries. This data analytical approach was developed in the early 2000s and has changed handling business data since then tremendously (Schermann et al. 2014). Simultaneously, the construction industry and, in particular, the as-performed construction stage provides a broad area for improvement, as many on-site processes haven't changed much in the last fifty years. The lack of structure causes tremendous cost and time uncertainties, which need to be resolved. Still today, the construction site represents a black box where crucial information gets lost, resulting in unexpected delays, high costs, and unsatisfied customers (Sacks et al. 2017).

The absence of development within the construction execution process leads to two significant problems. First, many construction projects are not executed as planned due to different reasons. For example, weather and bad on-site worker management can tremendously slow down the as-performed construction progress. Secondly, the construction domain has shown fewer efforts in precisely documenting the on-site progress.

The motivation for this research work can be explained in two parts: On the one hand, a well-designed method to acquire data automatically and seamlessly monitor construction sites needs to be introduced. On the other hand, a methodological approach to preprocess, interpret and understand the data is necessary to realize new developments for process optimization. There are hardly any studies on the as-performed construction process, which represent the implemented construction methods, processes, and resource inputs, which would enable an analysis of potentially avoidable delays and costs. The lack of recent developments in data from construction sites is closely related to the fact that less research has been done on automated monitoring techniques. This study intends to give a first approach to capturing various sensor data from construction sites and derive detailed information and knowledge in a second step.

A solid foundation to enable an analysis of construction processes is essential for this research work. Process modeling is a convenient technique to structure and document project behavior within construction projects (Cole 1991). Dori (2016) wrote that construction tasks have different dependencies upon each other that define the exact execution order. The End-

Start relationship, also called the precedence relationship, is the most common one in use. In this relationship, a task cannot be started until all of its predecessors have been completed. The construction tasks with their interrelationships can be visualized by a precedence graph representing the overall project. For now, this is mainly done manually, making it error-prone.

The main contribution of this paper is a workflow using data mining to get a step closer to automated construction monitoring. To do so, we performed the following steps: (1) Investigation and adoption of several data acquisition techniques on construction sites. (2) Creation of a data pipeline for transferring diverse data. (3) Processing the data and implementing ways to extract information from the sensor data. (4) Initialisation of data validation methods.

The individual stages of the introduced investigation of the as-performed construction process will be described in the following sections. Starting with a background section, all theoretical concepts and methods used in this paper will be explained. The second part of this paper introduces an approach embedding state-of-the-art concepts of the construction research domain with data mining approaches demonstrated by giving insights into implemented software and hardware components. Finally, the data collected from a construction site is analyzed, evaluated, and discussed further by giving ideas of industry-orientated application.

2 BACKGROUND

2.1 *Digital Twin Construction*

Sacks, Brilakis, Pikas, Xie, & Girolami (2020) developed in their research work on lean construction the concept of digital twin construction (DTC), which follows the plan-do-check-act cycle. This cycle of production control, introduced by Deming (1982), represents continuous improvement techniques of various workflows. Causes for delays, unexpected costs, and fraud can be detected early by construction managers on construction sites, currently appearing as black boxes.

The DTC process can be organized into four stages which are significant for creating an as-built model: Stage 1: Raw monitored data – Stage 2: As-built product – Stage 3: Project and process status knowledge – Stage 4: Simulation and analysis result.

To monitor the particular construction activities, specific monitoring techniques are necessary, and data pipelines must be created. In a second step, the data must be interpreted with complex data processing. Once extracted from raw data, the results give insights into the as-built product and the as-performed construction processes. It is essential to verify this data to ensure further data reliability. Therefore, project schedules, BIM models, and other construction-relevant data for verification should be used. The completion of this step ultimately leads to the third stage, which gives sufficient information about an ongoing construction project. The final step is to predict the performance of specific on-site construction activities, which can be summarized in the fourth stage by displaying results (Sacks et al. 2020).

The problem is that many scientists have thought about what needs to be done to create a meaningful digital twin considering construction sites, but not how (Opoku et al. 2021). This significant research gap yields applications for modern data mining techniques, eventually answering process performance questions. The employment of advanced data analysis techniques for analyzing the as-performed construction process, aiming to solve practical questions of the digital twin methods, is part of this research work.

Data mining has been an emerging technology in recent years and has proven its potential in other industries. We used the CRISP (Cross Industry Standard Process for Data Mining) (Cleve and Lämmel 2014), to apply DM to the construction execution.

2.2 *Analysis of the most refined process level*

To enable a data-based investigation of the as-performed construction stage, respective construction processes need to be broken down to the finest level, e.g., from the construction of the building to formwork placement. Wu et al. (2010) present a methodology to define fine-granular on-site processes, illustrated by a bridge construction example. The development of the process description is realized by adopting sequential Level of Details (LoD) on both the product and the process view. On the first level, a particular section of a construction project is selected. In the case of building construction, this can be *building A*. Then, the individual process model components of the corresponding sector are defined, which can be the *basement, first floor, second floor, and roof.* On the third level, more refined elements are specified, which are needed to build the construction parts mentioned before. This can be *walls, ceilings, columns, or beams*.

Finally, in the last step, a construction method can be selected for the particular construction elements, such as "cast-in-place". This results in the corresponding activities being generated representing the finest level of process description: *Place formwork, fill formwork, remove formwork.* Wu et al. (2010) used process patterns to formalize and encode the aforementioned construction methods in a computer-interpretable way. This represents the construction companies' knowledge of how to execute specific construction activities by combining process components and their precedence relationships. The method allows parallel construction while nevertheless enforcing that sub-processes must be finished before the next process can start. The process patterns include activity packages specifying materials, workers, and machines required to complete a low-level construction activity. Besides bridge projects, the process patterns have found application in shipbuilding (König et al. 2007) and building engineering (Beißert et al. 2007). Based on the activity package, monitored data, e.g., the location of a worker, can be integrated into individual as-performed construction processes.

2.3 *Previous research*

Previous research has applied data mining in some areas of the construction industry, yet the application during the construction execution stage proves to be small (Yan et al. 2020). Instead, most application of data mining methods in the construction industry was made in energy, building occupancy, cost management, material performance, safety management, textual knowledge discovery, framework establishment, and building design.

Some progress has been seen by showing the first attempts to develop the digital twin idea based on location-oriented work (Zhao et al. 2021). They introduced a BLE-based location-management system monitoring task progress from the worker's presence. Their goal was to detect low-level construction processes' start and finish times and evaluate them by the Construction Flow Index (CFI) (Sacks et al. 2017). However, it quickly becomes apparent that this research is still in its infancy while looking at the details of the case study. The construction site provided optimal conditions, as only renovation work was monitored. In addition, the approach could not distinguish between several persons, and a more precise distinction between smaller tasks was not possible. This undoubtedly shows that an automatic sensor-based construction monitoring task is not possible without further proceedings. More, the approach was demonstrated by only one case study, which yields the advantages of creating a digital twin framework. On the other hand, the monitoring method of Zhao et al. (2021) demonstrated investigation potential in construction planning and control by revealing an innovative way to unveil transparencies on construction sites. In conclusion, a more general data workflow must be developed to monitor construction activities on a wide range of different construction sites.

2.4 *Research gap*

Despite several developments in automated construction monitoring in recent years, it is unclear how to get information about the most significant parts of a construction site automatically. Moreover, previous case studies have failed to provide feasible solutions to interpret extensive sensor data from construction sites. To date, sensor-based monitoring approaches have only been applied to specific construction scenarios, which does not reflect general as-performed construction behavior. A major drawback of previous work is that the studies were only focused on one data source. However, different data sources need to be combined to generate the information, insight, and knowledge required.

In addition, more investigations are required in data processing to connect the gained results with the as-planned data and state-of-the-art site managing methods. Further, the thresholds of construction sites, e.g., project budget, sensor placement, and interference, are an intriguing area of the field of construction site monitoring that has not been considered sufficiently yet. Therefore, feasible approaches using different sensor systems according to the given thresholds of a construction site need to be developed.

This paper is a contribution to getting closer to automatic construction site monitoring. Our work aims to develop a feasible monitoring method that can investigate activities in specific areas, while combining different data sources, which may help us understand why delays happen on construction sites and give us additional information about the construction process. In addition, an approach to interpreting the raw sensor data and investigating single construction processes addressing activity packages is presented.

3 PROPOSED SOLUTION

3.1 *Framework for construction site documentation*

As a first approach to analyze the monitored data, we identify typical reasons for construction delays, as pointed out by Al-Momani (2000), like bad weather or suboptimal on-site conditions. This further helps to understand which data can be used for particular interpretations. To come up with an objective for the construction monitoring study and understand the construction progress, we define these initial hypotheses: (1) Construction speed is lower on construction sites that are exposed to stormy weather conditions. (2) A fail-safe access to material, labor, and equipment resources stabilizes the construction progress and reduces the overall cost. (3) The complexity of a building section strongly influences the construction behavior and, as such, the overall on-site performance.

The hypotheses mentioned above include but are not limited to aspects we want to investigate in this research work. However, we are not aiming to prove the introduced hypotheses in this paper. The reason for phrasing them is to establish a general direction of our research work. Following the data-information-knowledge pyramid, we want to derive information like construction speed, weather conditions, on-site resources, construction progress, and more from on-site monitored data. This will enable an information-based investigation of the as-performed construction process, building the foundation for business knowledge.

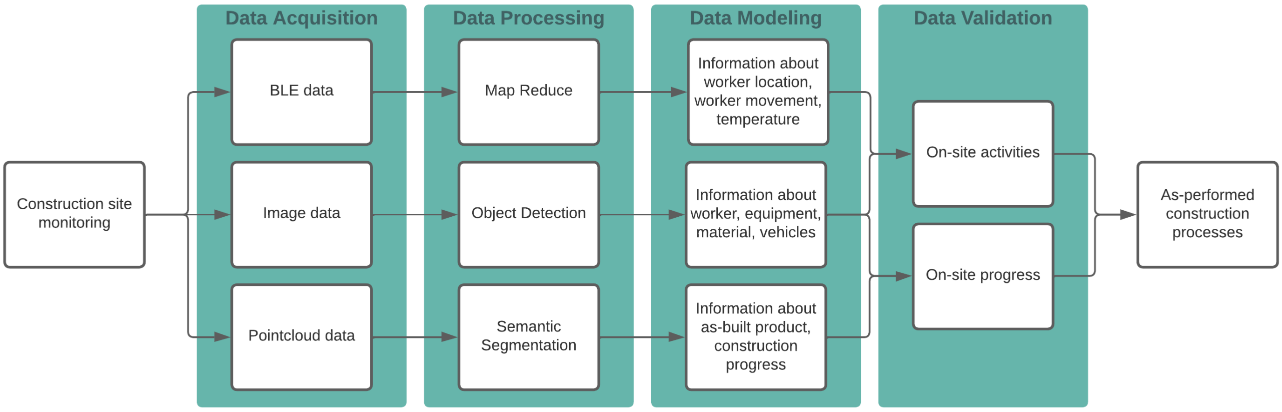

Figure 1: Big-data pipeline for on-site documentation

The chosen automated construction site documentation pipeline contains data mining methods and implements the concept of ”digital twin construction”. The approach consists of four essential stages, depicted in fig. 1.

3.2 *Data Acquisition*

The first stage is data acquisition. Different devices, like Bluetooth Low Energy (BLE) beacons, crane cameras, drones, and laser scanners, are used to get BLE, image, and point cloud data from construction sites.

Monitoring technique	Resulting data
Drone-camera	Images, photogrammetric point clouds
Crane-camera	Images, photogrammetric point clouds
BLE-system	RSSI, temperature and acceleration data
Laserscanner	Terrestrial point clouds

Table 1: On-site monitoring techniques

The reason for collecting BLE data is to accurately track the location and activities of construction workers. A BLE system comprises four key components: Beacon, gateway, IoT server, and database server. The beacons can measure different data like received signal strength indicator (RSSI), temperature, and acceleration. The data gets transferred via a BLE network provided by the gateways. The gateways are usually located at fixed places with known positions and receive the measured data from the beacons in a predefined time interval. All gateways are constantly connected to an IoT server, representing a communication bridge between the gateway and database server. The database server receives the data and archives it in a database. Working with IoT data, the use-case of the data tremendously defines the choice of database (Rautmare and Bhalerao 2016). Retrieving a significant amount of data from construction sites from different beacon types, the system must be flexible and scalable. The schemaless structure within a NoSQL database allows us to store the data of different beacon types in the same format without adding additional objects. NoSQL databases can be horizontally scaled using multiple servers. In contrast to the single node limitation of SQL databases, there are possibilities to extend storage using multiple nodes on NoSQL systems making this data schema a more practical choice for our big data approach.

The second source for monitoring the construction sites are pictures of the on-site progress. Therefore, cameras are placed on cranes or neighboring buildings, which allow good coverage of the construction site. These cameras are connected to a local server system and are constantly taking images. The local server uploads the images in a predefined time interval to remote nodes.

The third data type to document the construction progress particularly are point clouds. Point clouds can be created from laser scans or images. Creating semantically enriched point clouds with laser scanning is done by a three-step procedure, as shown by Tang et al. (2010): Data collection with dense point measurements of the facility, data preprocessing, and modeling with the BIM. Braun et al. (2015) define four steps to generate a point cloud using a photogrammetric approach: Image acquisition, orientation, image matching, and co-registration. A summary of the various monitoring techniques and the resulting data sources is shown in table 1.

3.3 *Data Processing*

A popular big data approach that can disassemble large databases into small segments is map-reduce (Dean & Ghemawat 2008). Usually, many queries are needed to gain information from enormous unstructured BLE data. In this case, map-reduce enables constructing large aggregations of queries. The approach will be used to process the extensive acquired BLE-data tailoring specific questions, which will provide further information about on-site construction activities: When did the construction worker pass specific points of interest? What was the temperature profile throughout the day? To which times were workers active?

Therefore, location, temperature, and acceleration data are needed. Acceleration and temperature are directly derived from processed from the databases. To calculate the distance from the RSSI value, a signal-strength distance estimation, introduced by Dong and Dargie (2012), is applied.

A powerful method to gain information from images is object detection. Multiple elements of construction sites, like construction workers, vehicles, and equipment, can be detected using this approach (Xuehui et al. 2021). The on-site environment is highly relevant to determining the individual construction activities (Torabi et al. 2022). For this purpose, it is required to train models to detect all construction process relevant information on images. To investigate working areas, an additional processing step is needed. The bounding boxes come with a 2D position on the image, which can be projected to a plan. This helps to specify the location of a worker in a known 2D environment. In summary, the image processing method can be broken down into three key stages: (1) Get images with workers at the relevant timespan, (2) Detect and locate workers on images using bounding boxes, (3) Project detected bounding box location to 2D-floorplan

To reconstruct the construction progress, it is necessary to create an as-built model based on raw point cloud data. A common method to develop as-built models is semantic segmentation (Tang et al. 2010). The scope of experiments performed within this paper lies in processing BLE and image data with map-reduce and object detection techniques covering data acquisition and data processing. Semantic segmentation will not be discussed within this paper. The steps after the latter phases are still under development and therefore mentioned only for the sake of completeness.

3.4 *Data Modeling and Validation*

Once the raw data from different data sources is further preprocessed and refined, it is necessary to represent the results in an understandable form. In addition, the as-built data must be aligned with the as-planned data. This requires a 4D BIM-Model, covering information about the specific construction activities and the construction schedule. Within this step, the different sensor information, like the location of workers, materials, and equipment, will be fused to document particular on-site workflows and the on-site progress.

After Data Modeling is complete and a representative digital twin is created, the deployment of gained information needs to be done. Therefore, a categorization is necessary: What are the most significant outcomes? How are they influencing construction site performance? What needs to be done to avoid particular delays? Experts in the construction industry are necessary to evaluate the processed information. Only after categorization is it possible to distinguish whether a particular construction process can be optimized or not. Ultimately a data platform is required, inevitably part of all steps, where construction process-related information is available for end-users.

4 EXPERIMENTS

To validate our approach, we monitored a building construction site in Regensburg, Germany, with crane cameras, drones, and laser scanners, as mentioned in section 3.2. The same monitoring process is currently applied at three additional building construction sites in the greater Munich area, generating large-scale datasets to further develop our concept and document the as-performed construction processes. We focused on processing BLE data and images to get the locations and movement of workers and the daily temperature. We assumed that a worker was close to a gateway in the two-meter radius.

4.1 *Construction site monitoring based on a BLE system*

The construction workers carried the BLE sensors attached randomly to helmets, facilitating anonymization. The beacons provide RSSI, acceleration, and temperature data. The iBeacon format was used to transfer the received BLE data. It has a JSON structure, including a timestamp, RSSI, and MAC address of the beacons. In every second, the gateways received data from the beacons.These messages get combined into gateway documents which are then sent to the IoT server, where they get processed into a human-readable form and then get forwarded to the database server. The communication happened via Message Queue Telemetry Transport (MQTT) requests, transferring the information in a JSON format. Once the gateway documents reach the database server, the information is stored on a MongoDB database in a BSON schema, adding the encoding values of acceleration and temperature. In detail, the gateway documents contain values like timestamp, the beacon's temperature, signal strength, acceleration coordinates, and the gateway's MAC address for the precise allocation. Each database entry refers to the specific gateway identified by the MAC address from which it was sent. The software framework containing all steps was implemented using Microsoft Visual C# in the .NET Framework 4.7.1 environment. In the first experiment on the construction site in Regensburg, we collected 90 Gigabytes of BLE data. A total of 31.000 gateway documents were assembled, resulting in 800 million beacon documents.

After collecting the data from the BLE system, the map-reduce approach was performed. First, all related or connected information to the individual on-site activities needed to be reduced to the scope of the investigation. The querying was performed based on specific time intervals and particular gateways representing the points of interest. The first step decreased the number of protocols and, as such also, the size of the databases. The second step of the map-reduce method mapped the data to the desired information, e.g., the signal strength, acceleration, and temperature. This process' reduce method reduced the total number of entries down according to a predefined threshold, e.g., worker close to a gateway. Considering the location of a worker group, it is essential how many beacons were close to particular areas to evaluate the position history at a later stage. The final output of the map-reduce procedure was a dictionary, linking the gateway id with the desired information. The method was performed multiple times to get detailed information about the location and movement of the workers and temperature data.

The processed data was eventually used to investigate the on-site activities. A way to get information is to average the results over the day and visualize them graphically. In this case, the reduced outputs of the map-reduce approach were visualized with the help of the Matplotlib library.

4.2 *Construction site monitoring based on crane cameras*

Three cameras were placed on the crane at different heights and angles. This allowed capturing the complete site project. The cameras took a picture each minute during the construction working times. In total, we collected an amount of 93.300 images within five months.

To extract information from the massive amount of construction site images, we conducted the following steps: First, we determined specific time intervals of the construction schedule where we want to investigate particular processes, e.g., formwork placement. In a second step, we used object detection based on pre-trained networks COCO dataset and TensorFlow to detect images with persons on it. Then the bottom location of the bounding boxes of the detected construction worker was saved. We derived floorplans from as-design information provided by the BIM model to build a 2D tracking environment. After that, we created a transformation matrix using OpenCV based on reference areas known on the image and the floorplan. Examples of such reference areas are staircases, elevator shafts, or cranes. Finally, we determined the positions of a worker group at a predefined time interval according to a 2D floorplan. An illustration of the performed approach is shown in fig. 2. In conclusion, the approach enabled detecting areas of construction work, similar to the BLE beacons, using a second data source.

Figure 2: Image processing method using object detection to process the location of workers

5 RESULTS

5.1 *Outcomes of data acquisition and processing of diverse sensor data*

The approach introduced in section 3 provides a concept to monitor construction sites using different data sources and process data accordingly. First, by looking at the RSSI data of the BLE sensors, we could detect specific working areas of construction workers. Figure 3 shows a clear tendency when specific work has been performed in which areas: Formwork jobs near the staircase were done in the selected time interval. In conclusion, various process patterns were detected within predefined time intervals and in areas of interest using the BLE method. This information is a foundation for the process patterns described in section 2.2 and the data modeling stage.

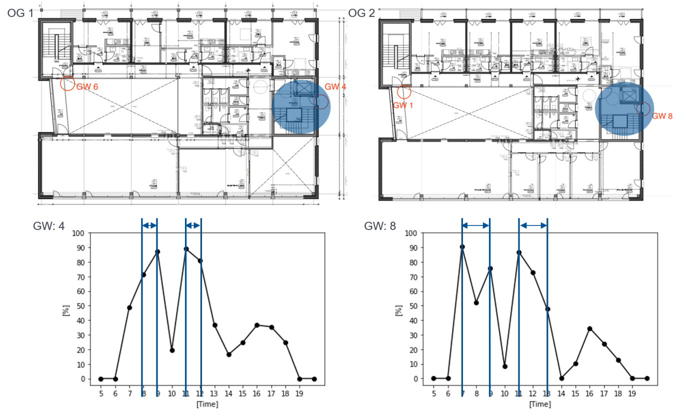

Figure 3: Working areas of interest

The on-site location and worker resources can be associated automatically. Until now, further experiments are planned to autonomously specify the building component, preconditions, and material resources.

Secondly, the BLE-acceleration data was used to detect the movement of workers during working times. Here we can identify the start and end times of the processes and breaks, fig. 4 shows a profile through the day.

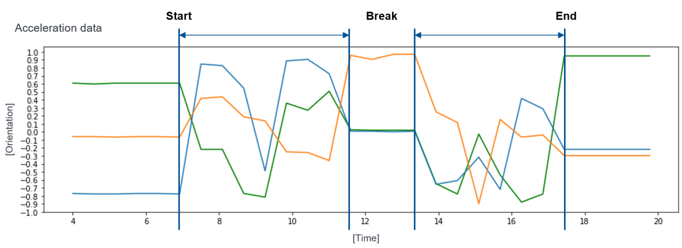

Figure 4: Start times, breaks and end times of on-site activities

Thirdly, to detect persons in a specific location with an additional data source, we conducted the object detection method based on images introduced in section 4.2. This allowed us getting an additional insight into working areas in predefined time intervals. More, this second data source specified the cases where the

BLE system was utterly wrong, improving the quality of the sensor dataset. Similarly, to the beacon system we could detect formwork work, shown in fig. 5, and various other process patterns.

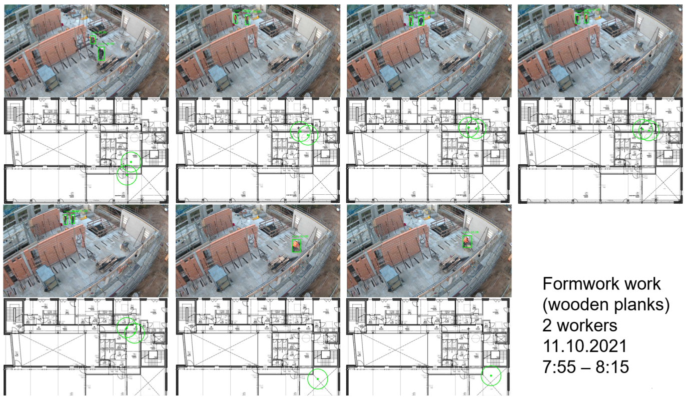

Figure 5: Results of image processing

Finally, temperature data was also measured and visualized. A temperature profile of a day is indispensable, especially when trying to find more about external influences on the construction process. Even though temperature data and movement data might not be directly found in the process patterns, explained in section 2.2, they allow further investigation of on-site happenings. This is highly important when considering the research objective introduced in section 3.1 since additional data sources, like weather data, our resources are required to understand the influences of the as-performed process and might give insight into causes of delays.

A fixed BLE-sensor was placed permanently close to the gateway to check if the proximity functions were working. The slight deviation, see fig. 6 due to standard BLE-interference can be considered the usual threshold and is not influential since we are not aiming for the most precise results. The general functionality (worker close to a specific point of interest) was fulfilled. In addition, the work areas detected with the BLE system and the work areas determined from the images of the crane cameras coincided.

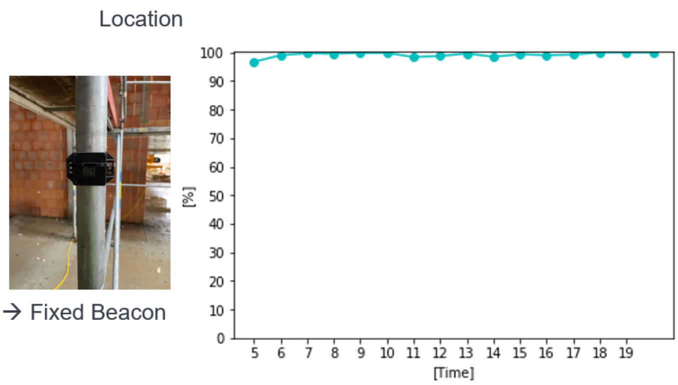

Figure 6: Fixed beacon verifying general functionality

5.2 *Discussion*

This paper introduced a conceptual framework for investigating the as-performed construction process based on data captured on-site by a multitude of different sensor system and its subsequent analysis and fusion. It investigated in detail the various data acquisition techniques for generating BLE data, images, and point clouds. On this basis, the paper discussed how big data processing approaches, such as map-reduce and image processing techniques, are employed to handle and analyse the large data sets. Within experimental setups of the sensors, we showcased that our data acquisition approach is feasible and can be implemented on real-world construction sites.

First, the BLE system proved to be a valuable data source for detecting the location and movement of workers. The placement of such a system is limited to the construction site restrictions resulting in uncovered areas. A data mining method, namely map-reduce, was applied to process vast amounts of BLE data enabling problem-oriented results, which were processed further to enable graphical demonstrations. The map-reduce method fused the data so that start- and end times of working days were detected, including specific working areas, movement, and temperature.

Second, a crane-based camera system was installed to take images of the on-site progress continuously. The images of the construction site proved to be beneficial for covering additional information about the as-performed construction process. The applied object detection method on the acquired images allowed precise localization of the construction works. We are further aiming to extend the image processing method to detect the movement of other equipment, like construction gear and vehicles, in order to document, when which equipment was used. At the same time, the movement data of the workers can be applied in a further step to encounter how much movement is necessary for which activity.

The concept behind collecting and processing data from diverse sources is to build a foundation for applying data mining technologies. As mentioned in section 3, this paper focuses on BLE and image data. Still, to monitor the as-performed progress, processing point cloud data, as realized for example by Braun et al. (2015), is indispensable. Findings and results regarding the integration with PCD will be published in upcoming papers.

Ultimately, the presented approach provides input for creating the digital twin of the construction site. However, more information is necessary to provide a detailed investigation of the as-performed construction process. Future work must extend the scope of data sources to build activity packages that can be compared to as-planned information. Once a solid foundation of process patterns is built, further investigation of the influences of the on-site process by receiving temperature data is possible. This requires more on-site data, which is still missing in the research domain, as also pointed out by others (Yan et al. 2020, Xuehui et al. 2021). In summary, the approach proved to be promising in detecting construction activities based on as-performed data, following the concepts of data mining and the DTC procedure towards enabling planned vs. actual comparison of on-site activities.

6 CONCLUSION

In this paper, we presented an approach to apply data mining techniques on construction processes by capturing and fusing diverse sensor data. We discussed a general concept on data capturing and step-wise processing and illustrated it by a number of highly relevant types of data sources, including images, point clouds and BLE-based location tracking. In addition, we demonstrated a method to fuse various data sources to create higher-level information and provide means for automated validation. We undertook this study to provide a reliable way to monitor and document construction sites without excessive manual work and required personal presence at the construction site. The approach demonstrated represents a further step in the direction of comparisons between as-planned models and as-performed sites. In the long run, the documentation based on various data sources should provide additional support for decision making in construction managing and allow future predictions of a project's performance. Undoubtedly, more research needs to be performed to realize full-scale automated on-site monitoring. This includes research on processing and fusing BLE data (Zhao et al. 2021), point cloud data (Xue et al. 2019), and image data (Torabi et al. 2022), concentrating on the workers' surroundings and considering the site environment. Construction diaries or reports merged with as-planned models are inevitable to extend the framework toward data modeling.

7 ACKNOWLEDGMENTS

The research presented in this paper was carried out with financial and operational support by Innovations Management Bau GmbH, the latter by providing us access to a number of construction sites. This support is gratefully acknowledged.

REFERENCES

Al-Momani, A. H. (2000). Construction delay: a quantitative analysis. *International Journal of Project Management 18*(1), 51–59.

Beißert, U., M. König, & H.-J. Bargstädt (2007). Constraint-based simulation of outfitting processes in building engineering. In *24th W78 Conference*, pp. 26–29.

Braun, A., S. Tuttas, A. Borrmann, & U. Stilla (2015). A concept for automated construction progress monitoring using bim-based geometric constraints and photogrammetric point clouds. *J. Inf. Technol. Constr. 20*(5), 68–79.

Cleve, J. & U. Lämmel (2014). *Data Mining*. De Gruyter Studium Series. Oldenbourg Wissenschaftsverl.

Cole, L. J. R. (1991). Construction scheduling: Principles, practices, and six case studies. *Journal of Construction Engineering and Management 117*(4), 579–588.

Dean, J. & S. Ghemawat (2008, jan). Mapreduce: Simplified data processing on large clusters. *Commun. ACM 51*(1), 107–113.

Deming, W. E. (1982). *Out of the Crisis*. MIT press.

Dong, Q. & W. Dargie (2012). Evaluation of the reliability of rssi for indoor localization. In *2012 International Conference on Wireless Communications in Underground and Confined Areas*, pp. 1–6. IEEE.

Dori, G. (2016). *Simulation-based methods for float time determination and schedule optimization for construction projects*. Ph. D. thesis, Technische Universität München.

König, M., U. Beißert, D. Steinhauer, & H.-J. Bargstädt (2007). Constraint-based simulation of outfitting processes in shipbuilding and civil engineering. In *Proceedings of the 6th EUROSIM Congress on Modeling and Simulation*. Citeseer.

Opoku, D.-G. J., S. Perera, R. Osei-Kyei, & M. Rashidi (2021). Digital twin application in the construction industry: A literature review. *Journal of Building Engineering*, 102726.

Rautmare, S. & D. M. Bhalerao (2016). Mysql and nosql database comparison for iot application. In *2016 IEEE International Conference on Advances in Computer Applications (ICACA)*, pp. 235–238.

Sacks, R., I. Brilakis, E. Pikas, H. S. Xie, & M. Girolami (2020). Construction with digital twin information systems. *Data-Centric Engineering 1*.

Sacks, R., O. Seppänen, V. Priven, & J. Savosnick (2017). Construction flow index: a metric of production flow quality in construction. *Construction Management and Economics 35*(1-2), 45–63.

Schermann, M., H. Hemsen, C. Buchmüller, T. Bitter, H. Krcmar, V. Markl, & T. Hoeren (2014). Big data. *Business & Information Systems Engineering 6*(5), 261–266.

Tang, P., D. Huber, B. Akinci, R. Lipman, & A. Lytle (2010). Automatic reconstruction of as-built building information models from laser-scanned point clouds: A review of related techniques. *Automation in construction 19*(7), 829–843.

Torabi, G., A. Hammad, & N. Bouguila (2022). Two-dimensional and three-dimensional cnn-based simultaneous detection and activity classification of construction workers. *Journal of Computing in Civil Engineering 36*(4), 04022009.

Wu, I.-C., A. Borrmann, U. Beißert, M. König, & E. Rank (2010). Bridge construction schedule generation with pattern-based construction methods and constraint-based simulation. *Advanced Engineering Informatics 24*(4), 379–388.

Xue, F., W. Lu, K. Chen, & A. Zetkulic (2019). From semantic segmentation to semantic registration: Derivative-free optimization–based approach for automatic generation of semantically rich as-built building information models from 3d point clouds. *Journal of Computing in Civil Engineering 33*(4), 04019024.

Xuehui, A., Z. Li, L. Zuguang, W. Chengzhi, L. Pengfei, & L. Zhiwei (2021). Dataset and benchmark for detecting moving objects in construction sites. *Automation in Construction 122*, 103482.

Yan, H., N. Yang, Y. Peng, & Y. Ren (2020). Data mining in the construction industry: Present status, opportunities, and future trends. *Automation in Construction 119*, 103331.

Zhao, J., E. Pikas, O. Seppänen, A. Peltokorpi, et al. (2021). Using real-time indoor resource positioning to track the progress of tasks in construction sites. *Frontiers in Built Environment 7*, 59.