

Digital Twin Enabled Asset Anomaly Detection for Building Facility Management

Xiang Xie*, Qiuchen Lu**, Ajith Kumar Parlikad***, Jennifer Mary Schooling****

**Institute for Manufacturing, University of Cambridge, Cambridge,
CB3 0FS UK (e-mail: xx809@cam.ac.uk)*

*** Bartlett School of Construction and Project Management, University College London,
London, WC1H 0QB, UK (e-mail: qiuchen.lu@ucl.ac.uk)*

**** Institute for Manufacturing, University of Cambridge, Cambridge,
CB3 0FS UK (e-mail: aknp2@cam.ac.uk)*

***** Centre for Smart Infrastructure and Construction, University of Cambridge,
Cambridge, CB3 0FA (email: jms33@cam.ac.uk)*

Abstract: Assets play a significant role in building utilities by undertaking the majority of their service functionalities. However, a comprehensive facility management solution that can help to monitor, detect, record and communicate asset anomalous issues is till nowhere to be found. The digital twin concept is gaining increasing popularity in architecture, engineering and construction/facility management (AEC/FM) sector, and a digital twin enabled asset condition monitoring and anomaly detection framework is proposed in this paper. A Bayesian change point detection methodology is tentatively embedded to reveal the suspicious asset anomalies in a real time manner. A demonstrator on cooling pumps is developed and implemented based on Centre for Digital Built Britain (CDBB) West Cambridge Digital Twin Pilot. The results demonstrate that supported by the data management capability provided by digital twin, the proposed framework realizes a continuous condition monitoring and anomaly detection for single asset, which contributes to efficient and automated asset monitoring in O&M management.

Keywords: Building Information Modeling, Digital Twin, Facility Management, Asset Management, Condition Monitoring, Anomaly Detection.

1. INTRODUCTION

The emerging of advanced computerisation has revolutionised the way building utilities are designed, constructed, operated and even maintained. Disruptive technologies, represented by Building Information Modelling (BIM), enable a dynamic, open access and digital built environment that achieves effective and efficient data storage, analysis, exchange and integration within the whole lifecycle of a building (Lu et al. 2019). More importantly, serving as an information source and a repository, BIM establishes a join between data across the building utility and three-dimensional geometric model. It is estimated that the adoption of BIM contributes to a reduction in time for updating databases in operation and maintenance (O&M) phase by 98% (Ding and Drogemuller 2009), owing to the capability of BIM digesting graphical and nongraphical information. As the main trigger that increases the whole lifecycle cost of a building, facility management (FM) represents an integrated approach to operate, maintain, improve and even adapt a building to promote a fertile environment that supports its strategic needs, such as maintaining the wellbeing and productivity of its occupants (Barrett and Baldry 2009). Achieving intelligent FM relies on comprehensive information, and various computerized maintenance management systems (CMMSs) have been developed to support FM tasks using information embedded in BIM (Kassem et al. 2015, Parn et al. 2017). In the

implementations, seamless information exchange is enabled between BIM and FM systems.

As the facilitator for FM efficiency, BIM is obviously not the ultimate panacea. Lu et al. (2019) summarized the limitations and gaps of BIM for achieving smart asset management (AM) in O&M phase/facility management (FM), from the perspective of technology, information, organization and standard. For instance, not all the FM related data is suitable to be hosted in a BIM environment, and automated data analysis tools for BIM hosted data have not been fully exploited and utilized yet. On the basis of BIM, the development of “Digital Built Britain” (DBB) is accelerating the adoption of digital techniques in the FM sector. As the report data for the public good (NIC 2017) stated, the UK needs a *Digital Framework for Infrastructure Data*, which could increase the performance, efficiency and resilience of the national infrastructure system. Hence, the derived digital twin (DT) concept is generated and promoted. DT is defined to be a realistic digital representation of physical assets, processes and systems, which could mimic their real-world behaviours. Actually, it has evolved into a comprehensive approach to manage, plan, predict and demonstrate building assets by encapsulating all the FM information required for efficient O&M through BIM data conduit. In order to maximize the value of DT and illustrate how DT support asset condition monitoring and anomaly detection in daily O&M

management, this paper presents a DT enabled asset condition monitoring and anomaly detection framework used in FM and the adopted structure of data integration.

Particularly, FM professionals are concerned about asset anomaly detection for various embedded systems in buildings, such as mechanical system, electrical system and plumbing system. Because assets are responsible for undertaking most of the service functionality of a building. Studies are conducted on asset monitoring, which can enhance the safety, efficiency and quality of building operation processes (Shi and O'Brien 2019), indicating that the combination of BIM and automated asset monitoring system contribute to the improvement of indoor environment quality and rapid intervention in case of hazardous conditions. Yu et al. (2014) reviewed analytical-based, knowledge-based, data-driven methods implemented for anomaly detection of air-handling unit. Most of the literatures for other assets follow similar classification scheme, and this paper demonstrates a functioning building FM system that enables automated asset condition monitoring and anomaly detection using data-driven method. Essentially, anomaly detection is a preventive and proactive action that guarantees assets maintain their anticipated function within lifecycle. A contextual anomaly detection, Bayesian on-line change point detection method (BOCPD), is adopted to detect where the generative parameters of the building operational data sequence drift, which is better than conventional point anomaly detection in theory.

2. BUILDING DT SYSTEM ARCHITECTURE

By definition, DT realizes a mean to link digital models and simulation engines with real-world data, which helps organizations to make better informed decisions, leading to improved outcomes. From a practical point of view, in addition to the information assimilation capability from BIM, four aspects of requirement are desired for the successful implementation of DT enabled FM, that are intelligence, efficiency, integration and interoperability. Intelligence allows for a transaction from conventional labour-intensive FM towards a predictive and automated one; efficiency requires the ability to manage assets in O&M with reasonable resource allocation; integration addresses the data compatibility problem and makes sure that data from different assets and processes could be integrated for further collaboration; while interoperability describes that DT enabled FM systems can be coherently dealt with various activities and seamlessly cooperated with other systems/people. This paper presents a typical hierarchical architecture for building DT that meets the requirements. It is comprised of five layers, which are data acquisition layer, transmission layer, digital modelling layer, data/model integration layer and service layer. Essentially, the ambition of implementing DT is to maximize the value of data, and harness data to create assets which are more responsive to requirements, perform better over their lifetime and create beneficial synergies across the wider built environment.

2.1 Multi-tier DT architecture

Layers for building DT architecture are presented in Fig. 1.

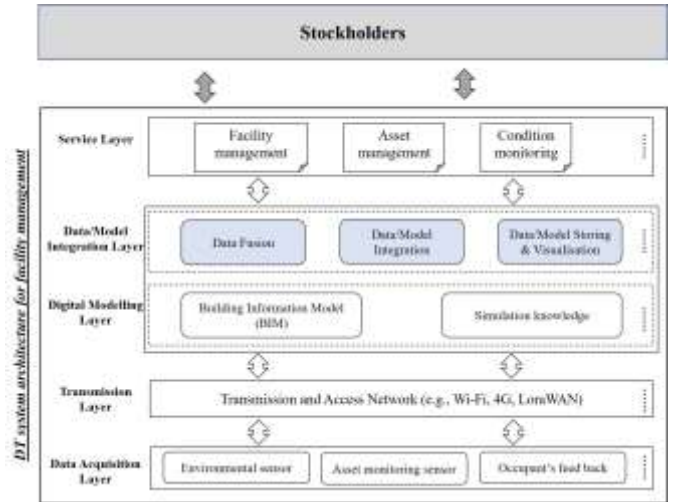


Fig. 1. System architecture for building DT.

The data acquisition layer is the foremost and the most fundamental layer for each DT. Although the objective of DT is to extensively collect data from different sources and maximize the value of data, data are not captured without costs. Even for the simple and cost-effective digital data, a certain amount of money is needed to purchase and maintain corresponding hardware. Therefore, in this layer, it is important to analyse the information/data requirements and list possible candidate techniques in data sensing. In regard to the building FM, few types of data are necessary. First, environmental data, such as temperature, humidity, air quality or light intensity, needs to be sensed for inferring the working status of assets according to their performance and evaluating the satisfaction level of occupants at the same time. Second, although there are tens of thousands of assets within one building, there are few principle assets with higher priority. Once these assets are not working properly, catastrophic consequences could be caused. For instance, if the boiler scales seriously, over-heating or even explosion may happen. Therefore, the principle assets need to be monitored in a real time manner. Internet of Things (IoT) sensors are gaining wide popularity because of cost-efficiency, low power consumption and elimination in maintenance need. Last but not least, occupants' feedbacks are the realistic reflection of the satisfaction of human beings. The role of feelings or experiences can never be replaced sensor reading or other technical metrics. Collecting opinions inside the spaces within buildings contributes to a better understanding about the effectiveness of FM.

The transmission layer aims at transferring data acquired to higher layers for modelling and analysis. Variety of communication protocols could be adopted according to the transmission range, data speed and data volume required in the specific scenario. Short-range communication networks include Wi-Fi, Zigbee and NFC, while long-range communication networks contain 4G/5G and LoraWAN (Centenaro et al. 2016).

Digital modelling layer is where BIM sits. In DT, BIM acts as a data conduit, offering great potential to retain and integrate

geometric and nongeometric attributes in a structured digital format (Cavka et al. 2017). Using BIM, any FM relevant information on 3-dimensional (3D) entity model, not only geometric data but also temporal parametric data throughout the lifecycle of assets which represents events and actions occurred over time, could be retrieved during the O&M phase. And it becomes much easier to manage, share and exchange comprehensive and incremental knowledge among different stockholders. In addition, simulation engines are also embedded in this layer. Energy and environment analysis, structure analysis, facility condition assessment and other building-level simulations heavily depend on the as-built or as-design drawings. Most of this information is stored in BIM.

Data/model integration layer is the core layer in the whole system architecture. By integrating all the data, simulation and domain knowledge through BIM data conduit, the as-is condition for the assets, processes and systems in building are analysed and updated. Live knowledge engines (KEs) are enabled with the capability to drive the simulation by assimilating data continuously and armed with intelligent functions (e.g. AI, machine learning modules), better-informed services are delivered.

Service layer is the top layer of the DT system architecture. It is formulated as a translator that interprets the knowledge and insights acquired from data/model integration layer to site-users and enables the interaction between stockholders and DT system. Most of the users of the DT systems and its derivative functions are not experts in building systems. They can only receive explicit suggestions in their decision-making process and give feedback back regarding their satisfaction for the decisions. The service layer realizes the isolation between users and technical details, which facilitates the usability of the DT system.

2.2 West Cambridge Digital Twin Demonstrator

The proposed DT system architecture is implemented in the West Cambridge site of the University of Cambridge. Specifically, the Institute for Manufacturing (IfM) building is used as a testbed for building-level DT demonstrator. This is a 3-story building, that includes teaching, studying, office, research and laboratory spaces and stands over a 40000-square-foot comprehensive area.

Based on the proposed DT system architecture, the IfM DT demonstrator is established integrating multiple data sources and providing multiple comprehensive applications, in order to satisfy different stockholders with different interests and requirements (shown in Fig. 2).

In regard to the data sources, the building management system (BMS), asset management system (AMS) and space management system (SMS) are integrated in the established DT through MySQL database. In addition to the embedded sensors and actuators in BMS, commercial IoT sensors are purchased and deployed in the IfM, as a supplemented data source. As discussed in the last section, two kinds of sensors are considered, including environmental sensors, which are used to evaluate the space comfort, and asset performance

sensors, which are utilized to measure the working condition of assets. Moreover, more than two hundred assets/spaces are tagged with QR codes, as an important way to collect information and feedbacks from the stockholders of the IfM, such as facility managers, site workers or occupants.



Fig. 2. The DT development of the IfM building.

Taking advantage of all the data available, several services are provided to the IfM building, such as ambient environment monitoring and maintenance prioritization and scheduling. In particular, real-time condition monitoring and anomaly detection for the variable refrigerant flow (VRF) cooling pump is investigated and implemented as a service in this paper.

3. ASSET ANOMALY DETECTION

Assets existing in every building, responsible for undertaking most of the service functionalities of the building sector, determine the quality of service that a building can provide to its occupants. Therefore, monitoring the working condition of the assets and further revealing the possible anomalies in a real time manner is widely investigated for optimizing building operation in FM. Specially, the detection of anomalies happened on a specific asset is challenging and problematic due to the high complexity of building systems and the large number of components in this highly integrated system. Two common strategies are usually adopted here to achieve real time condition monitoring and anomaly detection for assets. On one hand, for principle assets/components which undertake critical functionalities, sensors are usually attached to monitoring their working conditions through a few key performance indexes (KPIs), such as vibration frequency of pumps. Any significant variation in these KPIs indicates either a scheduled operation or a possible anomaly. By filtering out normal operation induced variations, candidate anomalies are recorded for further survey. Alternatively, considering that there are hundreds of thousands of assets in a commercial building, it is impractical to monitor every single asset using

at least one sensor. Therefore, a common way is to monitor the quality of service (QoS) provided by a system combining of multiple assets, such as HVAC system etc. For instance, by continuously logging the indoor spatial temperature/CO₂ concentration, the working condition of air handling unit could be inferred. Although it is extremely hard to acquire insightful understanding about the status of every componential asset according to the performance of the whole system, clues regarding the working condition of parts could be at least found. In this paper, the first strategy is investigated, aiming at proposing a general framework for realizing real time monitoring for principle assets in buildings.

When investigating the condition monitoring and anomaly detection solution for single asset, several approaches are widely adopted, including knowledge-based, analytical-based, and data-driven approaches (Andriamamonjy et al. 2018). Knowledge-based methods require prior knowledge of the working asset, from which simplified relationships or so-called rules can be extracted (Schein et al. 2006). Such procedure is simple to implement but often requires expert's knowledge, and the effectiveness of the procedure heavily depends on expert's qualification and experience. Analytical-based methods rely on first principle models that are typically based on physical laws such as mass and energy balance. By constructing a virtual analytical twin for the working asset, residual analysis or parameter estimation can be conducted to realize asset anomaly detection. Residual analysis simply evaluates the disparity between measured and simulated values, in which a beyond-threshold residual indicate anomaly (Wang et al. 2012). Alternatively, parameter estimation assumes that anomalies can be reflected with some specific parameters in the analytical twin. By identifying parameters in physical model, typical types of anomalies can be deduced (Bonvini et al. 2014). Reducing the dependence of data, analytical-based procedure can detect unknown anomalies effectively. However, it is found that researchers seem to lose their interests on this, because of the considerable efforts taken in implementing and calibrating the model.

Data-driven approach are promising solution because they completely depend on the past building operational data without requiring any physical interpretation about the systems (Chandola et al. 2009). Specifically, anomalies can be classified into two categories, that are point anomaly and contexture anomaly. If an individual data instance is diagnosed to deviate from its normal condition, the data instance is regarded as a point anomaly. On the other hand, if a data instance is anomalous under a specific context scenario, it is termed as a contextual anomaly. In the point anomaly detection, the normal operation conditions must be defined based on historical operation data, which serve as baselines and hereafter compared with current operation to detect anomalies. However, the operation loads to assets are changing throughout time, which causes the baseline behaviour fluctuating continuously. Considering that most of the existing methods are unable to handle the temporal contextual features of operational data, contextual anomaly detection analysis is widely studied to discover the association within datasets, where the external contextual attributes are used to reveal anomalous behaviour correlated with such

attributes. Change point detection is a form of contextual anomaly detection, which looks for significant variations or change points in the generative parameters of the building operational data sequence (Tartakovsky et al. 2012). The found change points are candidates for anomalies but not necessarily need to be an anomaly. Cross-referenced external contextual information must be integrated to support determine whether the detected change point attributes to the routine operation condition change or emerging anomalies. The workflow and information exchange behind the cross-referencing process are very complex. Fortunately, building DT happens to be the right solution that integrates multiple fragmented data sources and thus greatly enhances the data availability for buildings. Fig. 3 illustrate the block diagram of the asset anomaly detection framework.

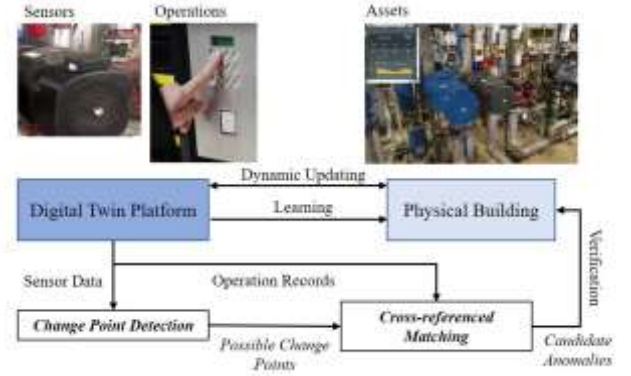


Fig. 3. Overview of the Building Asset Anomaly Detection Framework.

Here, a Bayesian Online Change Point Detection algorithm (BOCPD) is adopted (Kiasi et al. 2013), due to the fact that it can be used without requiring any contexture prior knowledge about the normal or anomalous conditions. Let $\mathbf{x} = \{x_1, x_2, \dots, x_N\}$ be an inputted sequence of observed anomaly indicative data from a monitoring cycle. The target of the adopted BOCPD based approach is concerned with estimating the posterior distribution over the run length at each time instant, in other words, the time interval since the last change point given the data observed so far. Let r_t denote the length of the current run at time t , which implies that the last change point occurs at time $t - r_t$ and the set of observed data associated with the current run is $\mathbf{x}_t^r = \{x_{t-r_t+1}, \dots, x_t\}$. Therefore, the targeted posterior distribution of the run length r_t can be formulated as:

$$p(r_t | \mathbf{x}_{1:t}) = \frac{p(r_t, \mathbf{x}_{1:t})}{p(\mathbf{x}_{1:t})} \quad (1)$$

where the observation likelihood is obtained by marginalizing over the joint likelihood of run length and observation: $p(\mathbf{x}_{1:t}) = \sum p(r_t, \mathbf{x}_{1:t})$. The run length is truncated to zero if a change point is encountered, otherwise the run length increases by one. Therefore, the joint likelihood can be updated online using a recursive message passing scheme:

$$\begin{aligned}
p(r_t, x_{1:t}) &= \sum_{r_{t-1}} p(r_t, r_{t-1}, x_{1:t}) \\
&= \sum_{r_{t-1}} p(r_t, x_t | r_{t-1}, x_{1:t-1}) p(r_{t-1}, x_{1:t-1}) \quad (2) \\
&= \sum_{r_{t-1}} p(r_t | r_{t-1}) p(x_t | r_{t-1}, x_t^r) p(r_{t-1}, x_{1:t-1})
\end{aligned}$$

Suppose that the interval between neighbouring change points follows a Geometric distribution with a constant expectation of the mean run length, and the underlying predictive model $p(x_t | r_{t-1}, x_t^r)$ complies with a Normal-Inverse-Gamma prior on independent identically distributed Gaussian observations, which is computationally advantageous. Eventually, the change points with maximum posterior probability are considered as the potential anomalous events.

4. DEMONSTRATION AND VERIFICATION

In this section, the application of the proposed framework on monitoring two centrifugal pumps is described as a case. The two pumps of equal characteristics are installed in the plant room of IfM building. They work in parallel to pump return chilled water from the air handling units & fan coil units back to the chiller. For centrifugal pumps, typical faults like defective bearing could result in catastrophic consequences, such as abnormal noises, rotating unbalance, even shaft breakage. The most revealing and widely acceptable diagnostic information of centrifugal pump is the vibration measurements. Because the vibrations are transferred from pump outwards through its casing, for the convenience of measurement, featured vibration frequency measured by the sensor mounted at the pump casing close to the bearing is adopted as an indirect method of assessing the conditions inside the monitored pumps.

Supported by the IfM DT platform, a long period of vibration frequency data is available, which makes it possible to continuously conduct a tentative diagnosis for the pump health condition. The data include the response to both scheduled operating condition changes as well as a real fault resulting in strong abnormal noises. Two datasets with a sampling time of one hour are picked to examine the performance of BOCPD. In the first case, the studied centrifugal pump 1 undergoes a scheduled shutdown due to the bank holiday. The period of data starts from the 5th December of 2018 and lasts until 1st January of 2019 (4 weeks).

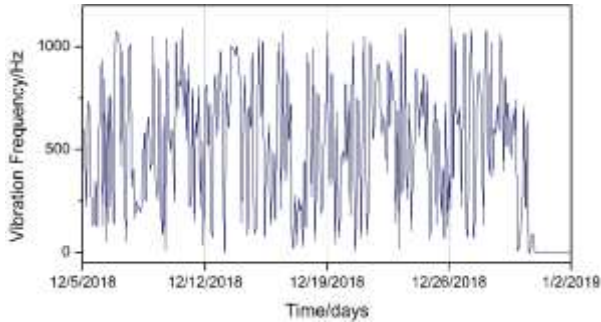


Fig. 4. Vibration frequency time series in the pump shutdown scenario.

Fig. 4. shows the recorded vibration frequency time series within given period. The shutdown can be seen to the naked eyes, and a rough judgement can be made that the studied pump stops working from the afternoon of 31st December of 2018. Fig. 5 depicts the output of the BOCPD based approach when applied to the pump shutdown event. In the figure, red solid line reveals the local maximum a posteriori run length estimation result, while blue dashed line marks the most probable run length considering the continuity of the run length. It can be found that although the local optimal run length appears some spikes, the subsequent data is more than enough to compensate the side effects caused by occasional measurement errors. This is the key point to the reduction of possible false alarms.

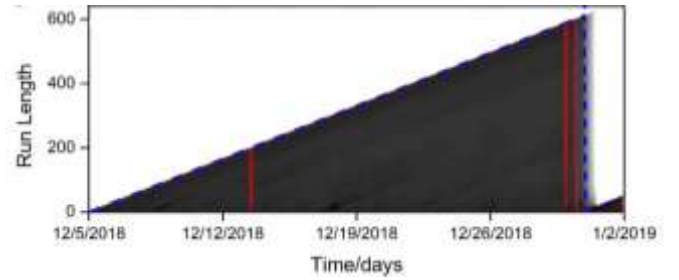


Fig. 5. Detection of the shutdown event by BOCPD procedure.

In the second case, one of the two pumps undergo a highly suspicious anomaly causing a strong anomalous degree of noise, while the other one works properly. Here, an artificial data set is generated by combining 14 days vibration frequency data from the normal pump with 14 days data from the faulty pump (from 9th July to 22nd July in 2018). Fig. 6 shows the generated vibration frequency time series within given period. Different from the shutdown scenario, it is extremely hard to distinguish the difference between the vibration of normal and faulty pumps by unaided eyes. Therefore, BOCPD is utilized to detect the change point between two kinds of vibration frequencies.

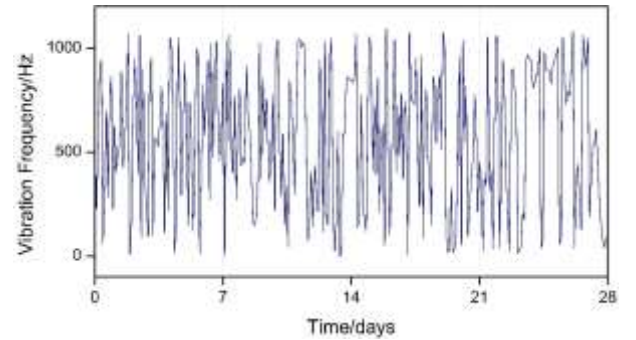


Fig. 6. Vibration frequency time series in the pump anomalous noise scenario.

Figure 7 depicts the output of the BOCPD based approach when applied to the pump fault event. Obviously, the BOCPD procedure shows a favourable capability of detecting changes with limited time delays. The red cross labels the fault

awareness time. The advantage of BOCPD is that despite there is delay before the faults of pumps are recognized, real change point time can be uniquely recognized when subsequent data is available. For the cross-over match process, a more precise change point contributes to the matching between symptoms and corresponding normal operation changes.

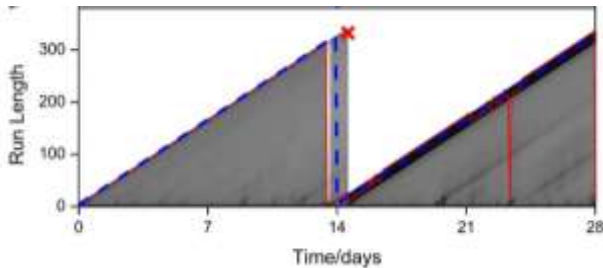


Fig. 7. Detection of the pump anomaly event by BOCPD procedure.

5. CONCLUSIONS

FM professionals attaches great importance to the real time asset condition monitoring and anomaly detection, because assets undertake vital roles in providing services to the occupants. To realize data-driven asset anomaly detection, building DT system is incorporated to integrate multiple fragmented data sources and enhance the data management capability and data availability for buildings. The detection framework is divided into two sub-tasks: implement Bayesian online change point detection, aiming at finding the time instant at which the underlying symptom parameters of sequential sensing data are suspected to change, and then distinguishing change points caused by routine operations and real anomalies. The proposed framework is verified on Centre for Digital Built Britain (CDBB) west Cambridge Digital Twin demonstrator, using centrifugal pumps as a case study. Results indicate that supported by DT, the proposed framework realizes a continuous anomaly detection for single asset.

ACKNOWLEDGEMENT

The research that contributed to this paper was funded by the EPSRC/Innovate UK Centre for Smart Infrastructure and Construction and the Centre for Digital Built Britain at the University of Cambridge.

REFERENCES

- Andriamamonjy, A., Saelens, D. and Klein, R., 2018. An auto-deployed model-based fault detection and diagnosis approach for Air Handling Units using BIM and Modelica. *Automation in Construction*, 96, 508-526.
- Barrett, P. and Baldry, D., 2009. *Facilities management: Towards best practice*. John Wiley & Sons.
- Bonvini, M., Sohn, M.D., Granderson, J., Wetter, M. and Piette, M.A., 2014. Robust on-line fault detection diagnosis for HVAC components based on nonlinear state estimation techniques. *Applied Energy*, 124, 156-166.
- Cavka, H.B., Staub-French, S. and Poirier, E.A., 2017. Developing owner information requirements for BIM-enabled project delivery and asset management. *Automation in construction*, 83, 169-183.
- Centenaro, M., Vangelista, L., Zanella, A. and Zorzi, M., 2016. Long-range communications in unlicensed bands: The rising stars in the IoT and smart city scenarios. *IEEE Wireless Communications*, 23(5), 60-67.
- Chandola, V., Banerjee, A. and Kumar, V., 2009. Anomaly detection: A survey. *ACM computing surveys (CSUR)*, 41(3), 15.
- Ding, L. and Drogemuller, R., 2009. Towards sustainable facilities management. In *Technology, Design and Process Innovation in the Built Environment*. Spon Press, 399-418.
- Kassem, M., Kelly, G., Dawood, N., Serginson, M. and Lockley, S., 2015. BIM in facilities management applications: a case study of a large university complex. *Built Environment Project and Asset Management*, 5(3), 261-277.
- Kiasi, F., Prakash, J., Patwardhan, S. and Shah, S.L., 2013. A unified framework for fault detection and isolation of sensor and actuator biases in linear time invariant systems using marginalized likelihood ratio test with uniform priors. *Journal of Process Control*, 23(9), 1350-1361.
- Lu, Q., Xie, X., Heaton, J., Parlikad, A.K. and Schooling, J., 2019, October. From BIM Towards Digital Twin: Strategy and Future Development for Smart Asset Management. In *International Workshop on Service Orientation in Holonic and Multi-Agent Manufacturing*. Springer, Cham, 392-404.
- NIC (National Infrastructure Commission), 2017. Data for the Public Good. See <https://www.nic.org.uk/wp-content/uploads/Data-for-the-Public-Good-NIC-Report.pdf>.
- Parn, E.A., Edwards, D.J. and Sing, M.C., 2017. The building information modelling trajectory in facilities management: A review. *Automation in construction*, 75, 45-55.
- Schein, J., Bushby, S.T., Castro, N.S. and House, J.M., 2006. A rule-based fault detection method for air handling units. *Energy and buildings*, 38(12), 1485-1492.
- Shi, Z. and O'Brien, W., 2019. Development and implementation of automated fault detection and diagnostics for building systems: A review. *Automation in Construction*, 104, 215-229.
- Tartakovsky, A.G., Polunchenko, A.S. and Sokolov, G., 2012. Efficient computer network anomaly detection by changepoint detection methods. *IEEE Journal of Selected Topics in Signal Processing*, 7(1), 4-11.
- Wang, H., Chen, Y., Chan, C.W., Qin, J. and Wang, J., 2012. Online model-based fault detection and diagnosis strategy for VAV air handling units. *Energy and Buildings*, 55, 252-263.
- Yu, Y., Woradechjumnroen, D. and Yu, D., 2014. A review of fault detection and diagnosis methodologies on air-handling units. *Energy and Buildings*, 82, 550-562.