A Digital Twin-Enabled Anomaly Detection System for Asset Monitoring in Operation and Maintenance

Qiuchen Lu [a]; Xiang Xie [b,*]; Ajith Kumar Parlikad [b]; Jennifer Mary Schooling [c]

a. The Bartlett School of Construction and Project Management, University College London, UK

b. Institute for Manufacturing, University of Cambridge, UK

c. Centre for Smart Infrastructure and Construction, University of Cambridge, UK

Abstract

Assets play a significant role in delivering the functionality and serviceability of the building sector. However, there is a lack of efficient strategies and comprehensive approaches for managing assets and their associated data that can help to monitor, detect, record, and communicate operation and maintenance (O&M) issues. With the importance of Digital Twin (DT) concepts being proved in the architecture, engineering, construction and facility management (AEC/FM) sectors, a DT-enabled anomaly detection system for asset monitoring and its data integration method based on extended industry foundation classes (IFC) in daily O&M management are provided in this study. Following the designed IFC-based data structure, a set of monitoring data that carries diagnostic information on the operational condition of assets can be extracted from building DTs firstly. Considering that assets run under changing loads determined by human demands, a Bayesian change point detection methodology that handles the contextual features of operational data is adopted to identify and filter contextural anomalies through cross-referencing with external operation information. Using the centrifugal pumps in the heating, ventilation and air-cooling (HVAC) system as a case study, the results indicate and prove that the developed novel DT-based anomaly detection process flow realizes a continuous anomaly detection of pumps, which contributes to efficient and automated asset monitoring in O&M. Finally, future challenges and opportunities using dynamic DTs for O&M purposes are discussed.

Keywords: Digital twin, Anomaly detection, Industry Foundation Classes (IFC), Operation and Maintenance management

* Corresponding author
E-mail addresses: xx809@cam.ac.uk (X. Xie)

1. Introduction

The Operation and Maintenance (O&M) phase for building and civil infrastructure assets covers more than 50 years of the total life span [1]. Achieving smart building management is a complex issue in the O&M phase. Comprehensive information needs to be recorded (e.g., historical O&M records, performances of facilities, accurate locations etc.) and multiple technologies would be involved (e.g., sensors, cameras etc.). Keeping data integrity, validity and interoperability is the key challenge during the process of O&M management [2]. Consequently, an effective and intelligent O&M management system is needed to maintain dynamic information, support various activities and contribute to a satisfactory environment [3]. Various tools and systems have been developed to improve O&M management, such as Computerized Maintenance Management Systems (CMMS), Computer-Aided Facility Management (CAFM) systems, Building Automation Systems (BAS), and Integrated Workplace Management Systems (IWMS) [4]. For instance, CMMS is a computerized system for O&M management, which can record daily work orders, historical records, service requests and maintenance information. But it still requires significant effort and time for facilities management (FM) professionals to extract the diverse O&M information they need (e.g., data within CMMS, specifications, 3D models) [2]. There is a lack of an integrated platform that could manage information distributed in different databases and support various activities in O&M phases. Advances in building information modelling (BIM) is likely to aid in reducing the time for updating databases in O&M phases by 98% [5]. Some integrated and comprehensive solutions for O&M management have been proposed by adapting BIM and developing systems to improve data interoperability and integration. For instance, Motawa and Almarshad proposed a Case-Based Reasoning (CBR)-integrated BIM system for building maintenance to improve the efficiency of decision making and communication among different stakeholders [6]. The restoration team of the Sydney Opera House also designed a unified central data repository integrating different resources to support effective O&M management. But overall, a comprehensive and effective data integration/query approach based on BIM, which can be maintained and updated throughout the O&M phase is still under investigation [5,7]. In summary, an integrated intelligent approach or system that can help to monitor, update, communicate and integrate O&M management issues is still required for continuous development and improvement.

During the O&M phase, anomaly detection for building assets, such as mechanical, electrical and plumbing systems (MEP), is considered not only the most labour-intensive and timeconsuming but also the most influential process [8]. Extensive studies demonstrate that timely anomaly detection could ensures the safety, efficiency, and quality of the building operation processes to a large extent [8]. Essentially, it is a preventive and proactive action that guarantees the assets maintaining their original anticipated function within their lifecycle. However, one of the big challenges is that these assets run under changing loads determined by human demands. Therefore their performance, for instance the pump vibration in the daily O&M, is not stationary. Conventional point-based anomaly detection algorithm cannot cope well with this, especially in the targeted built environments where the unavailability of well-labelled data is typical. In response to this situation, contextual anomaly detection, represented by Bayesian on-line change point detection method (BOCPD), becomes a promising alternative. Instead of anomalous points, change points are detected where the generative parameters of the building operational data sequence drift. Combined with the external building operation information, real anomalies that result in asset failures could be filtered as the trigger for following-up early warnings. Generally, the anomaly detection of asset monitoring for O&M management requires cross-referencing of multiple data sources for building facilities information. A comprehensive solution is necessary for streamlining anomaly detection, in which data interoperability and reusability need to be significantly enhanced.

Digital Twins (DTs) are considered to be such a comprehensive solution [9]. The concept of DTs evolved as a comprehensive approach to manage, plan, predict and demonstrate building/infrastructure or city assets. The DT is a digital model, which is a dynamic representation of an asset and mimics its real-world behaviour [10,11]. Moreover, due to the data analytical and decision-making capability DT possessed, the way we plan, deliver, operate, maintain and manage the assets is reinvented, thus better services can be provided [12]. To maximise the value of DTs and further present how they may support anomaly detection in daily O&M management, this study presents a DT-based anomaly detection system and an appropriate method of data integration based on the extended IFC. Then, a novel Bayesian change point detection methodology is adopted to indicate the suspicious anomalies of pumps, based on the building DT. This system is brought to life through the development of a dynamic demonstrator based on the West Cambridge Digital Twin Pilot.

2. Literature Review

2.1 Current Research on Daily O&M Management

Many existing O&M management approaches already benefit from emerging data capture and management technologies, for instance, radio frequency identification (RFID) [13], sensor systems [14,15,16], image-based techniques [17] or virtual reality (VR)/augmented reality (AR) [17,18]. As shown in Fig.1, technologies used in current O&M management can be classified as software, hardware, and network technologies.

Commonly adopted software tools include: computer-aided design (CAD), IWMS [23], CMMS, BEMS, BAS and enterprise asset management (EAM) [24], which can be used to manage daily activities and provide required services. A pilot construction project at the University of Southern California aimed at linking BAS, CMMS and Document Management Systems (DMS) with BIM and provide a demonstrator of BIM-to-BIM-FM in practice [25,26]. Due to the proliferation of a multitude of software supporting the different O&M and FM activities, accessing the required information can become difficult for FM professionals especially when information is stored in disparate systems. Hardware consists of equipment used in office and on-site (shown in Fig.1). Sensors and tags are gaining popularity in O&M to aid in the creation of a 'dynamic' and 'intelligent' asset management environment. Tags (e.g., QR code, RFID) and sensors connect scattered assets into an integrated unit, and further support real-time data collection and storage [27,28,29,30]. Network (i.e., web-based) technology can provide remote connections to different data resources and cloud-based services for different platforms [31].

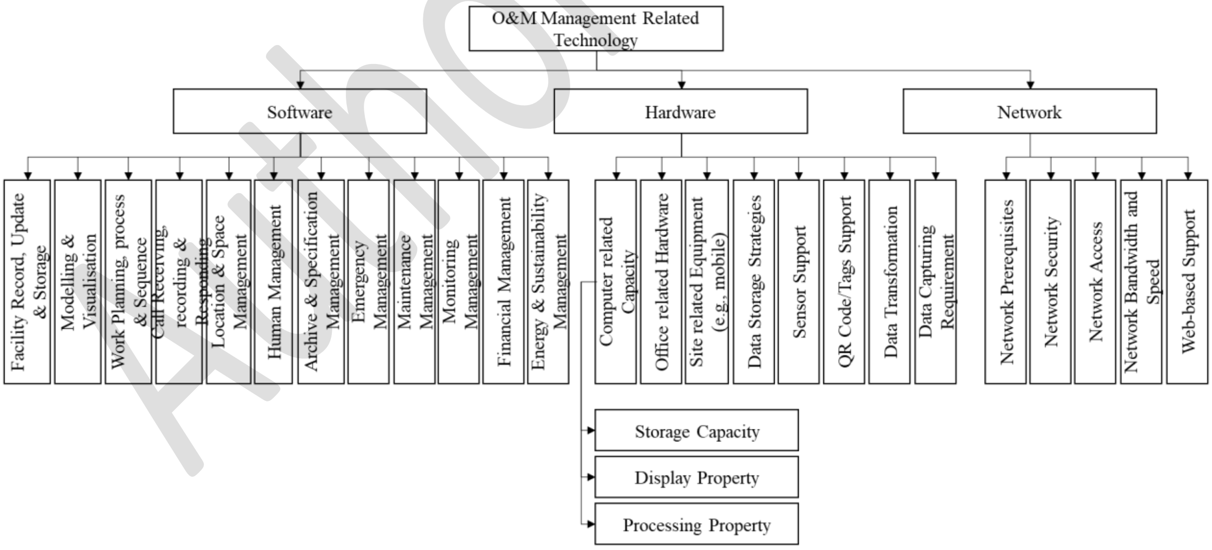

Figure 1 The functional descriptions of technology requirements for O&M management [19,20,21,22]

Alongside technologies, information directly determines the result of decision making in O&M [3,32]. Complex information (e.g., historical O&M records, space information, accurate locations etc.) is recorded and exchanged during O&M management processes (Fig.2).

Effective decisions usually depend on comprehensive, continuous, reliable and accurate datasets (e.g., asset information, as-is conditions) [39,40]. Hence, the integrity, validity and interoperability of information are crucial for improving management efficiency and intelligence [3,32]. The information required for O&M can be classified and listed as shown in Fig.2. Nongeometric information (e.g., building/infrastructure asset related information) can be directly integrated with geometric information via digital devices in the BIM environment. BIM-enabled asset management would further provide ease of access for information retrieval. Various practical studies and academic research have proved that BIM-enabled asset management provides long-term and obvious benefits [31,41,42,43,44]. The time and resources required in accessing relevant equipment and building materials information could be reduced [43]. For instance, Hassanain et al. [45] proposed an effective IFC-based data model for integrating maintenance management information. However, their work mainly focused on developing a generic framework and only used for roof objects. Hence, an appropriate method of data integration is still needed to further ease and benefit O&M information exchange and sharing.

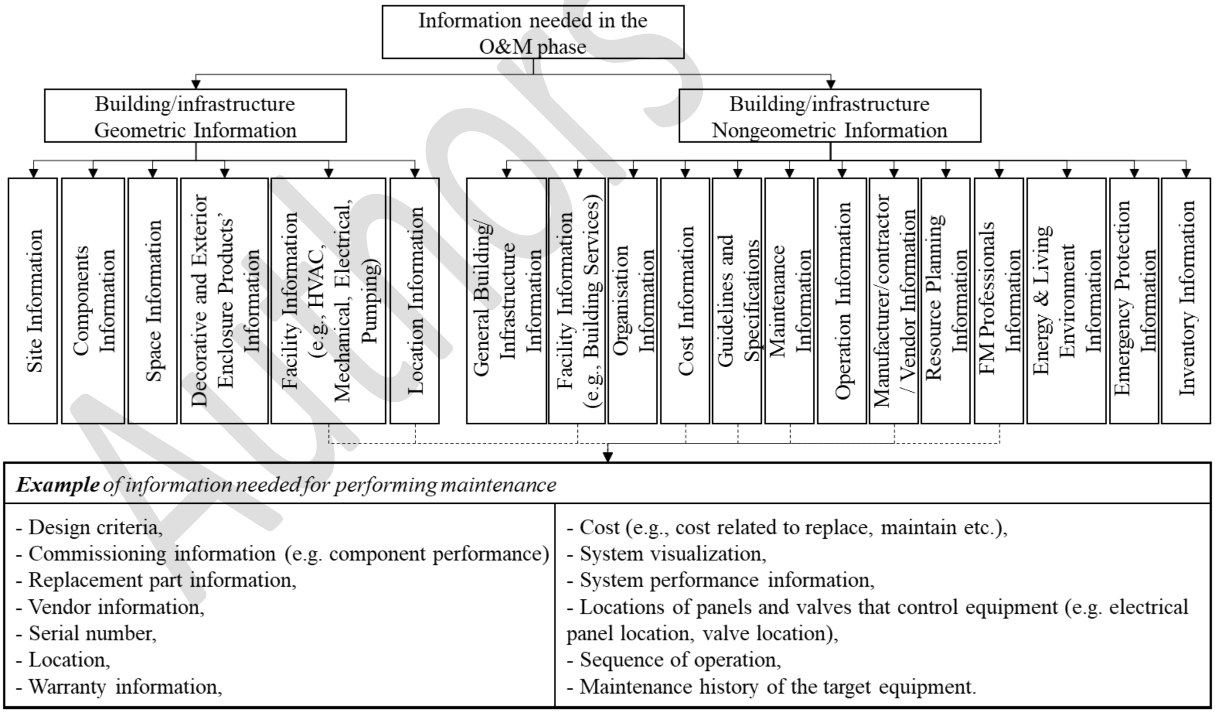

Figure 2 Information requirements in O&M phases [3,33,34,35,36,37,38]

Although a large amount of effort has been made in achieving smart O&M management, a lack of well-organised framework/system to link all assets efficiently, as well as the capability to manage required information, is one of the key problems in O&M management.

2.2 Review of anomaly detection techniques in buildings

Assets within the building, responsible for delivering the service functionalities of the building, determine the quality of service that a building provides to its occupants. Therefore, monitoring the working condition of the assets and further revealing the raised anomalies in a timely manner is widely investigated for optimizing building operations in the O&M phase. In particular, the detection of anomalies for asset monitoring is challenging and problematic due to the high degree of system complexity and large scale and the number of components in this highly integrated system. A common practice is detecting whether the performance of assets exhibit anomalies that deviate from the anticipated behaviours [46].

Specifically, anomaly patterns can be classified into two categories: point anomalies and contextual anomalies. If an individual data instance is diagnosed to deviate from its normal status, 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. For the mainstream point anomaly detection, the so-called normal operation conditions must be defined based on either historical operation data or model simulations, which serve as baselines and are thereafter compared with current behaviour to detect anomalies. Typically, process history-based methods are extensively adopted because they depend on the past building operational data without requiring any physical interpretation of the systems. Moreover, the data-driven nature makes these methods extremely easy and inexpensive to implement, as long as data satisfying quality requirements are available. For instance, Capozzoli et al. [47] adopt artificial neural ensembling networks to capture the dynamics behind the normal building energy consumption data. GESD many outliers detection algorithm [48,49] is used to analyse the dynamics residuals, identifying patterns of anomalies occurring in a cluster of buildings. Similarly, Magoules et al. [50] demonstrate the effectiveness of recursive deterministic perceptron (RDP) neural network in detecting anomalies in building energy consumption profiles. These methods assume that well-labelled data under normal operating conditions is available.

However, in practice, it is difficult to distinguish normal and abnormal operating conditions, which depends heavily on human evaluation for now. Therefore, the unsupervised anomaly detection techniques can be used to model the intrinsic property of the normal and abnormal datasets given limited prior knowledge, so that anomalies can be uniquely identified. Clustering techniques [51,52], such as hierarchical agglomerative clustering or entropy-weighted k-means (EWKM) method, are used to find anomalous behaviour in building energy data. The advanced quantitative association rule mining (QARM) is another promising technique [53,54,55,56], which is adapted to discover useful knowledge and derive rules from the unlabelled operational data. The rules discovered are used to identify raised anomalies. It is reported that these unsupervised techniques are useful in anomaly detection and operation pattern recognition for building assets [57].

The operating conditions and working loads on building assets are changing throughout time, which causes continuous baseline behaviour fluctuation. Considering that most existing methods are unable to handle the temporal contextual features of operational data, contextual anomaly detection analysis is 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 abrupt variations or change points in the generative parameters of the building operational data sequence [58]. More precisely, the found change points could be suspicious candidates for anomalies but not necessarily need to be an anomaly, serving as an early warning symptom for the problem within the underlying building system. For instance, Touzani et al. [59] adopt a statistical change point algorithm to detect potential “non-routine events” in building energy data, which provides a tractable starting point that can be expanded for discovering changes in operational characteristics and possible anomalies in building systems. Cross-referenced external contextual information must be integrated to help determine whether the detected change point attributes to the normal condition variations or emerging anomalies. However, the workflow and information exchange behind the cross-referencing process is very complex. Fortunately, DT of buildings is a solution that integrates multiple fragmented data sources and thus greatly enhances the data availability for buildings [60]. With the help of the DT model, normal operating condition changes could be excluded, leaving only the suspicious anomalies that help facility managers identify the problems as early as possible.

3. DT-based Anomaly Detection Process Flow

The process flows under two different scenarios (i.e., DT-based and traditional) have been established based on literature review [3,6,9,61,64], and expert interviews (i.e., facility management and estate management teams in authors’ university). Compared to the DT-based anomaly detection process, the traditional process shows two main defects, namely scattered information and manual query processes [3,6,9,61,64].

Even though some maintenance and operation data are managed in some facility information systems (e.g., BMS, AMS in Fig.3 and 4), it still requires a significant amount of time to search, query, verify and analyse the corresponding facility information from heterogeneous data sources. For instance, based on the expert interviews, data lists of each system have been summarised in Fig.3. When the FM professionals receive a maintenance request through the call service system (Fig.4), they need to search relevant information of the failed asset saved in the asset management system (such as historical information or manufacturer) first, and then confirm the location information saved in the space management system. If further required, some additional information may also need to be queried from BMS or other systems. Moreover, this process might also cause errors and deviations. The duplication of information queries frequently occurs in the traditional process. For instance, overlapping data may also be saved in different databases (e.g., historical records, locations and corresponding contractors' information) [3]. As shown in Fig.3, data sets of sites, buildings and floors are redundantly and repetitively saved in some systems, including AMS, BMS and SMS. Besides the scattered information, manual query processes are also the key problem of anomaly detection delay. In the traditional process, the facility manager usually acts as a central coordinator and their decision-making would depend on related information, as well as expert experience [6], as shown in Fig.4.

These problems of the traditional process indicate that there is a need for an intelligent and comprehensive platform to integrate and effectively search information, facilitate decision making and semi-automate/automate processes. In that way, with the consideration for the convenience of searching, verifying, querying and managing facility information and automating anomaly detection through a DT-based system, these problems can be improved and further addressed.

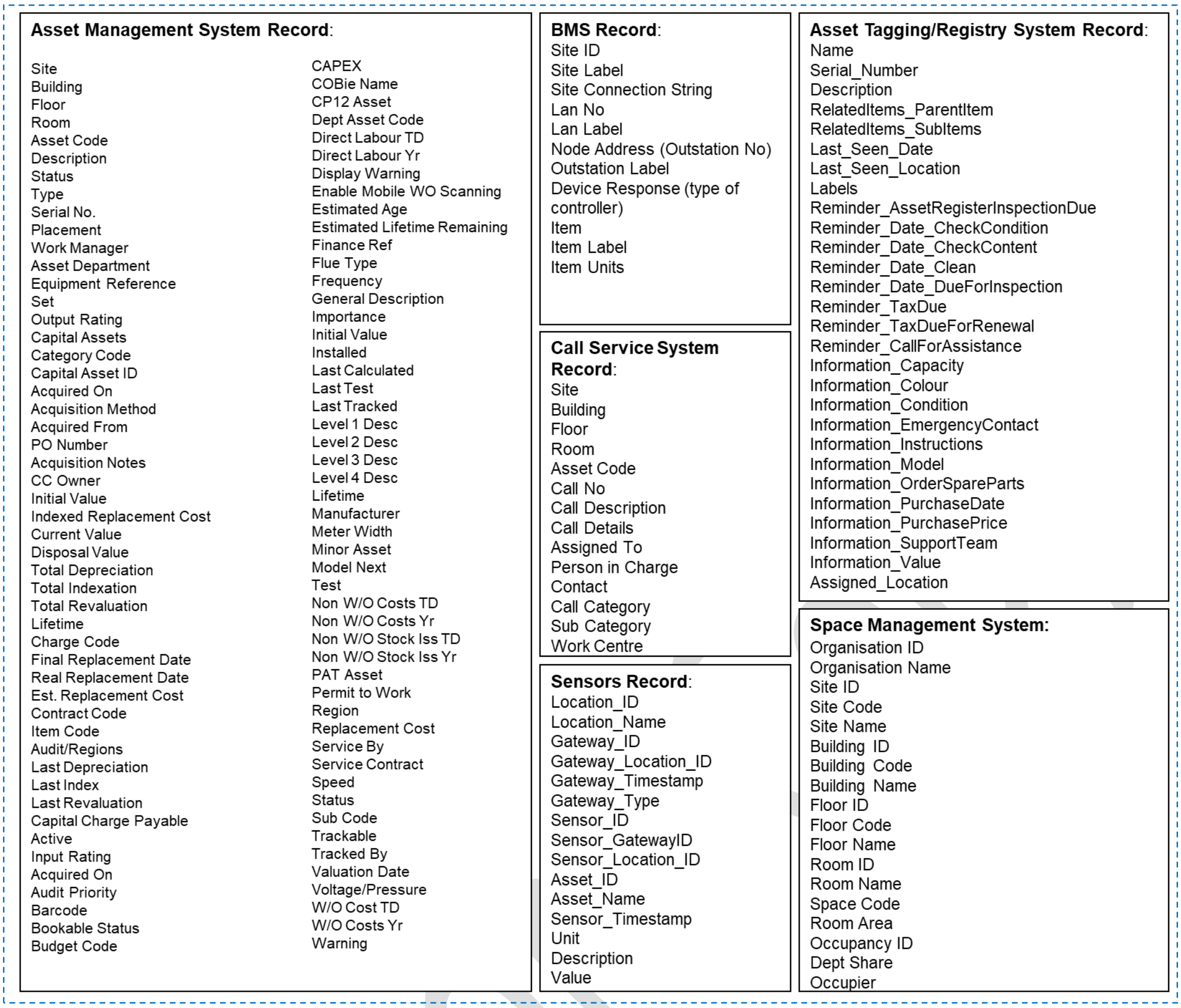

Figure 3 Data lists of each system in daily O&M management

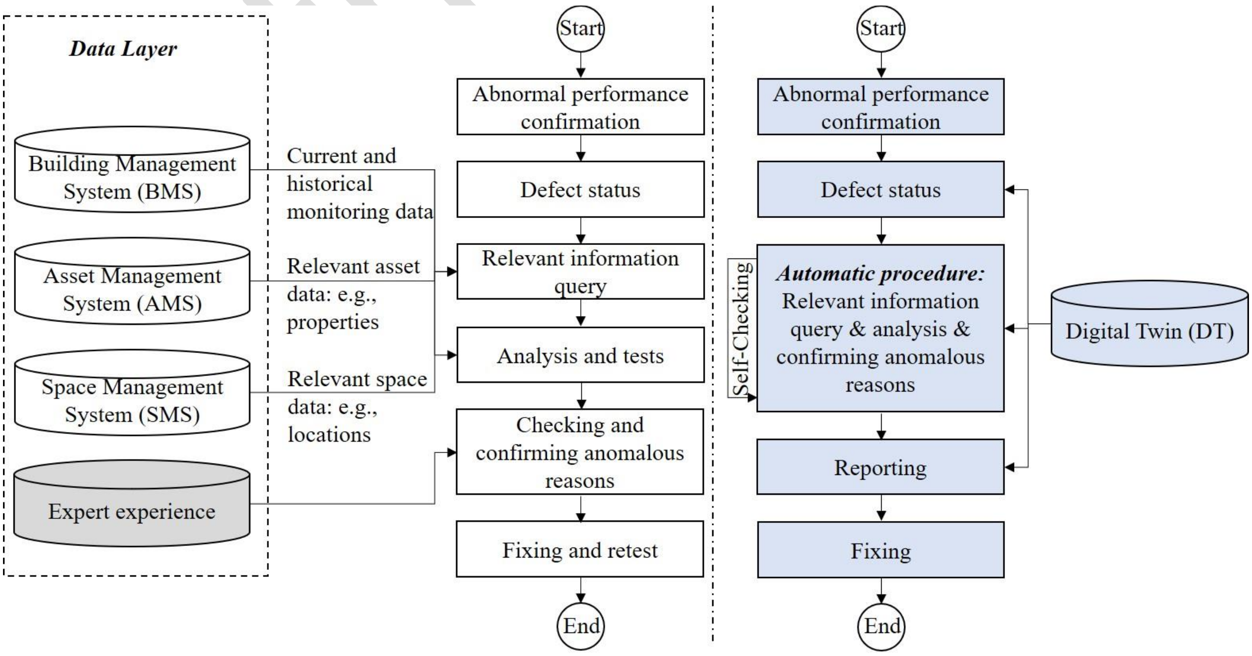

Figure 4 Anomaly detection process flows in O&M phases: scenario 1 (left) traditional process and scenario 2 (right) DT-based process

4. The DT-based Anomaly Detection Framework

4.1 Anomaly detection oriented data availability in existing buildings

Detecting anomalies of building assets in the O&M phase involves multi-domain and multi-layer information storage, manipulation, exchange and interaction. Effective data integration through information sharing is a critical factor in achieving effective anomaly detection, especially for excluding change points caused by normal operating condition changes, to avoid any false alarms. In addition to those commonly adopted tools (e.g. BAS, CMMS) introduced in section 2.1, anomaly detection in building O&M research also relies on other relevant data sources, such as the emerging sensing systems, access control systems or security cameras in buildings. Under the well-established communication protocols of building data storage and exchange, new data sources in O&M are still emerging. For a building HVAC system, the BAS data emerging from sensors and actuators (which might be Building Management Systems (BMS) in other cases) could be used federatively to detect the anomalous operating behaviour in a timely manner [62]. For instance, when the sudden drop in the supply air temperature of an AHU in heating mode is diagnosed, building sensing data (or access control system and security camera for occupancy monitoring in other cases) should be integrated to determine whether the drop is caused by an extreme change of outdoor temperature. However, if the supply temperature drops below its mixed air temperature, chances are that a potential anomaly happens in the AHU heating coil valve. The CMMS database keeps a detailed record of the occupants' service requests and work-order issues to address these service requests [63]. The inspection and maintenance data of CMMS could provide an insightful clue to enrich the building knowledge, like fault trees and relationships between components. Field expert rules can be acquired to enable the root-cause identification capability for possible anomalies in a building. However, the fragmented nature of building data sources presents a challenge in developing a valid anomaly detection strategy. The next section describes the DT solution provided to integrate multiple data sources that can support the anomaly detection task.

4.2 DT construction and data integration

Building DTs in this study were constructed based on definitions, namely 'DTs integrate their sub-DTs and intelligent functions (e.g., AI, machine learning, data analytics etc.) to create digital models that are able to learn and update from multiple sources, and to represent and predict the current and future condition of their physical counterparts correspondingly and timely' [64]. The DT's construction also follows the designed architecture provided by authors, referring to [9] and [64]. It includes five layers: data acquisition layer, transmission layer, digital modelling layer, data/model integration layer and service layer.

In practice, several O&M platforms and databases are used in daily management (e.g., BMS, SMS mentioned in section 3). The O&M data is usually saved in different formats. It thus requires great efforts and time for FM staff to extract the diverse and scattered O&M information required. A unified and standardised data schema is needed for information integration and achieving smart asset management in the O&M phase. Because of the flexibility and consistency of IFC schema in the building lifecycle, IFC schema is the most suitable and fundamental data schema for wider BIM implementation and information integration. Hence, the extension of the current IFC to fulfil O&M management requirements would be a critical step. Moreover, the asset information generated in the O&M phase is not static. For instance, sensor data is dynamic in real time and maintenance events would also be recorded case by case. A single IFC file would be ineffective for decision making and also difficult for additional information query, since existing IFC files may only include basic geometry information. Therefore, a possible and effective solution for representing IFC schema and integrating information is to provide a centralised data model linking with distributed data resources in daily O&M management.

Hence, in the data/model integration layer of building DTs, the data structure is designed to be capable of interchanging and interoperating external data related to each BIM object in the digital model on a semantic level, to enable IFC-based interoperability between BIM and other data sources. The IFC is used as the central data model and other data resources are kept in their original storage locations, which are saved in this distributed manner.

All the current research provides solid evidence of the increasing attention of BIM development in FM. However, research that systematically studies IFC in O&M phases is missing. There are no entities in the existing IFC4 schema to specifically represent information and activities in O&M phases [20]. With these considerations, more subclass data entities, types and parameters required for FM should be extended for DT data structure development. More complicated data types and specific O&M activities need to be provided [34,65]. Data schema about the inspection and maintenance process needs to be defined, and omitted properties and relationships related to FM need to be supplemented [39,62,65].

To update the O&M information to as-is DTs and map the data model of maintenance and inspection activities into the IFC standard, IFC extensions are proposed and developed based on the maintenance and inspection activities, required information and process as the core step of DT construction. In this research, IFC4 is used as the base specification for introducing new entities. In IFC4 schema, *IfcProcess* can present the activity or process of an activity/event/task/procedure for a building project. It usually happens in building construction with the intent of designing, costing, acquiring, constructing, or maintaining products [66,67]. However, the maintenance and inspection processes are required to be included in IFC schema, including inspection events, maintenance events and required actions/resources. *IfcControl* is the abstract generalization of control or constraint products/processes in general, which covers the specification, regulation, cost schedule or other requirements [66,67]. Even if *IfcControl* can represent the partial required information about the maintenance plan, schedule and cost, these entities are not initially designed for O&M management and thus cannot be completely matched with O&M activities. *IfcActor* defines a person or organization involved in a project during its life cycle. Specific roles in the O&M phase are not well defined and classified. *IfcRoleEnum* only includes one role type about FM, namely FM manager. *IfcAsset* presents an identifiable grouping of elements with financial values. However, more information is required in FM, for instance, history record and status of assets (as shown in Table 1). Moreover, specific asset types should be developed and classified for O&M management. For instance, *IfcAssetTypeEnum* should be further designed for FM and *IfcCostItem* needs more items to be added related to O&M management. *IfcAsset* needs to be extended for the O&M phase.

[**Insert**: Table 1. Evaluation of IFC4 support for O&M management information requirements]

In addition, one of the most important information records in the O&M phase is the historical record of the asset, but neither *IfcOwnerHistory* nor *IfcPerformaceHistory* cover complete information relevant to FM. For instance, there is no enum designed for FM in *IfcChangeActionEnum*. Table 1 lists the details of how asset register requirements can be matched with IFC4 entities and COBie 2.4 spreadsheet. Some requirements cannot be directly linked with entities in IFC4. Most of these unmatched data are important elements during O&M phase, including lacking capital information (e.g., costs breakdown, source of components and spare parts, and consumption) and incomplete information (e.g., history record, maintenance cost, and maintenance activities) (as shown in Table 1).

COBie is one of Information Exchange national standards (in the US, UK, and other countries) successfully adapted in the industry and the most relevant IE specification that can be implemented for the integration between BIM and O&M systems. On the other hand, partial information required for O&M can be presented using the COBie.Job worksheet [68], or FM software can provide the information manually/semi-automatically through ad-hoc functions. However, COBie is still immature from some technical perspectives: 1). model validation after the information exchange is needed; 2). user-friendly information save and query approaches and formats are required; 3). clear classification strategies of assets in O&M phases (e.g., sensors and control points) are needed to avoid misunderstanding of various O&M activities.

Assets in O&M phases can be classified into service-related assets and monitor & control-related assets according to their functions and relationships with existing buildings (Fig.5). Service-related assets (e.g., HVAC systems, lighting systems etc.) provide daily O&M services and refer to specific assets belonging to parts of existing buildings. Monitor & control-related assets (e.g., sensors) are additional assets attached to existing buildings/systems and equipped with monitoring and controlling functions. As shown in Table 1, subclass entities need to be included in the existing schema. The entity *IfcProcess* and the entity *IfcControl* are suggested to be extended and two corresponding subclass entities (*IfcOperationandMaintenaceProcess* and *IfcOperationandMaintenanceControl*) can be added to represent the maintenance and inspection activities. *IfcAsset* should be further extended for FM based on O&M requirements. Three subclass IFC entities are also suggested to be developed for enhancing O&M information management, namely *IfcMaintenanceHistory*, *IfcInspectionHistory* and *IfcSpareRecord* (Fig.6).

The data integration method provided in this research integrates information in a distributed and dynamic way. Based on the primary IFC file, required additional IFC entities are first added to the existing IFC files. Then, the matching tables for other database integration are created for describing the relationship between the BIM object GUID and its corresponding database ID from other data sources (e.g., AMS). When relevant data (saved in AMS) needs to be integrated or queried for some services in the DTs, the matching table provides a linking bridge between the targeted BIM object (GUID) and the corresponding ID in other data sources (e.g., AMS) (as shown in Fig.7). In this way, this data integration method enables that IFC and other data sources (e.g., AMS) are independent of each other, while keeping linkages. Thus, all data sources (including BIM, AMS etc.) can be updated individually and kept dynamically.

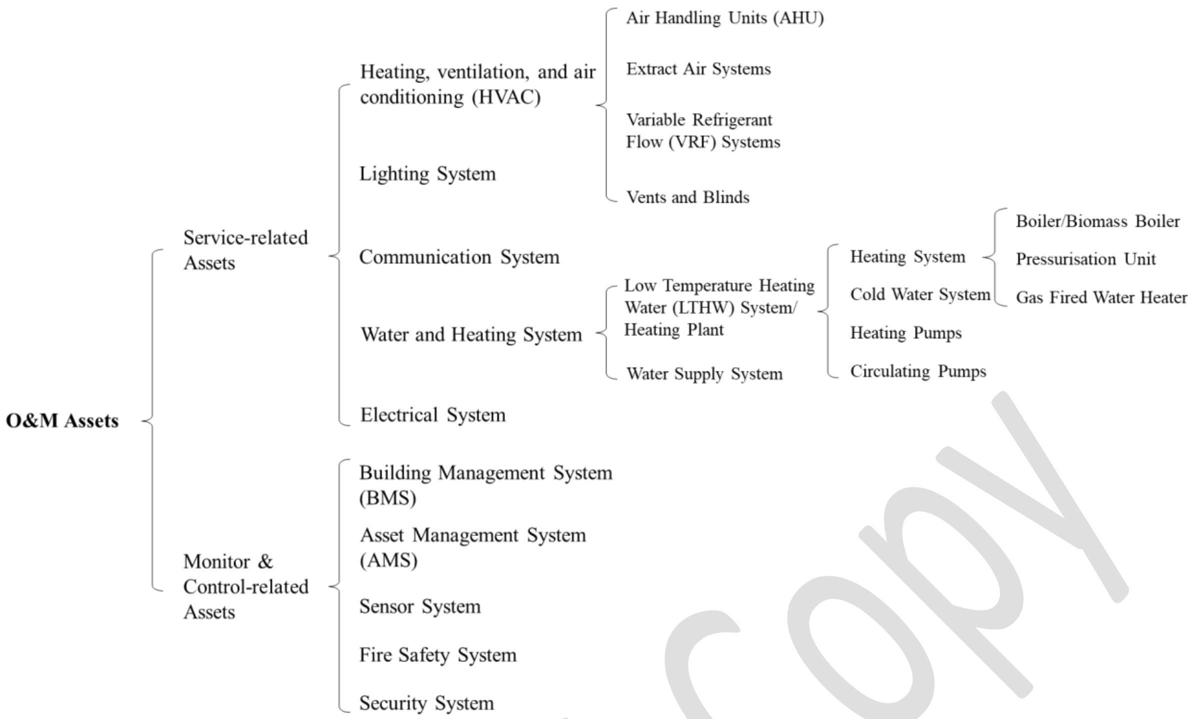

Figure 5 Asset classification in the O&M phase

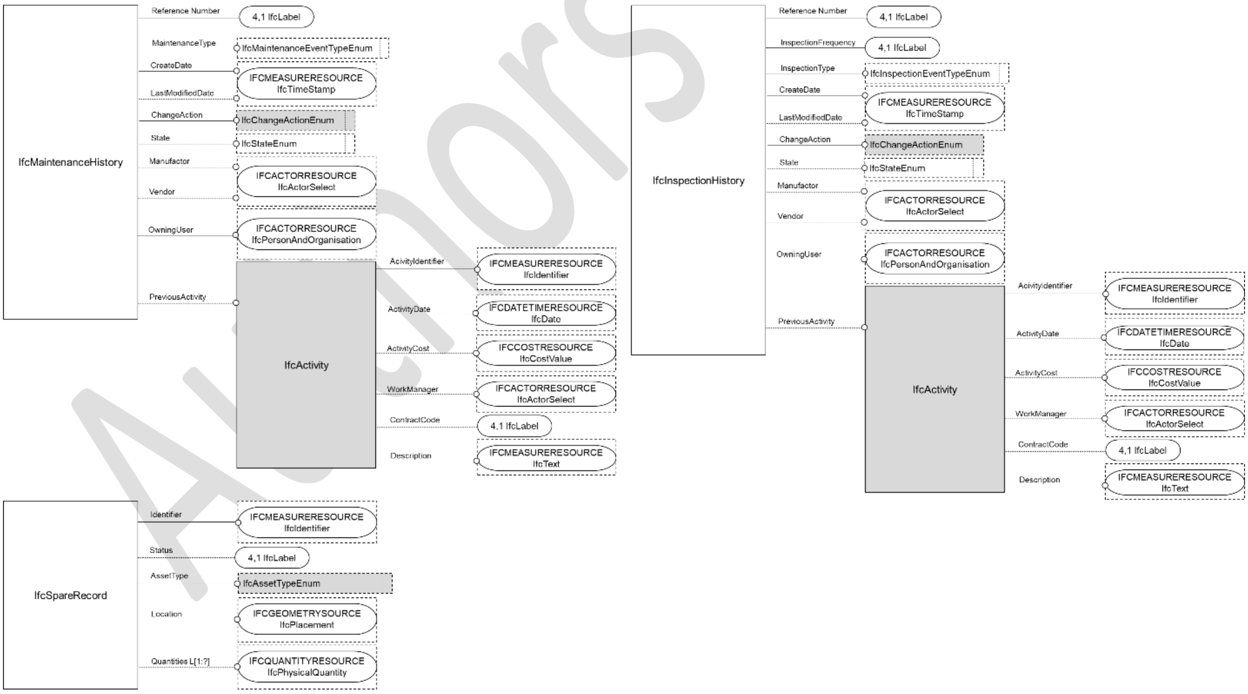

Figure 6 Three suggested entities (the grey blocks present additional properties)

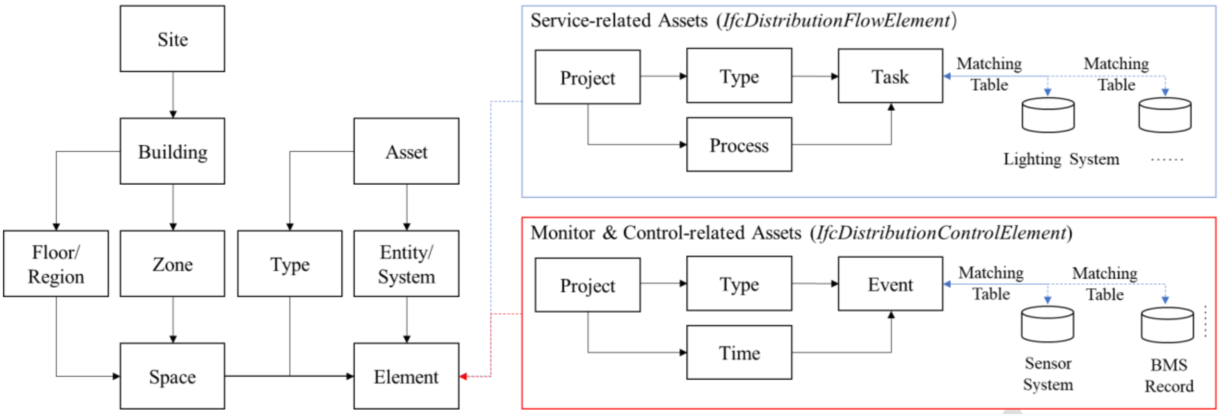

Figure 7 Asset management and record between BIM and other data sources

4.3 Anomaly detection procedure for asset monitoring

In this section, a general procedure for asset anomaly detection is illustrated to implement the monitoring of asset anomalies using data managed with the IFC schema that carries diagnostic information on the operational condition of assets. A block diagram of the framework can be seen in Fig.8. The whole procedure is divided into two sub-tasks: (1) change point detection, aiming at finding the time instants at which the underlying symptomatic parameters of sequential operational data are suspected to change, due to either operating condition variations or emerging anomalies; (2) anomaly identification, aiming at distinguishing change points caused by logged operating condition variations or real anomalies through event matching. For change point detection, different from most of the statistic methods, such as cumulative sum or likelihood ratio test [69], the BOCPD [71] is a natural approach to segment sequential data and can be used for online anomaly detection without requiring prespecified thresholds, which are difficult to establish a priori. Upon finding change points in operational data, a simple cross-over matching is conducted to identify change points caused by actual anomalies, thus eliminating the points resulting from normal operation condition variations and keeping the false-alarm rate to the minimum. Generally, the BMS (which might be BAS in other cases) keeps detailed track of the building system operational processes. Therefore, we could simply consider that change points identified around recorded operational variation time are normal reactions, while other unclaimed change points are the consequence of suspicious anomalies on corresponding assets. Afterward, appropriate responses can be provided promptly. Since the cross-over match process is quite instinctive, this paper focuses on the change point detection algorithm. It is also worthwhile noticing that this procedure is general, thus it can also be implemented on assets in any building system, such as HVAC system and MEP system.

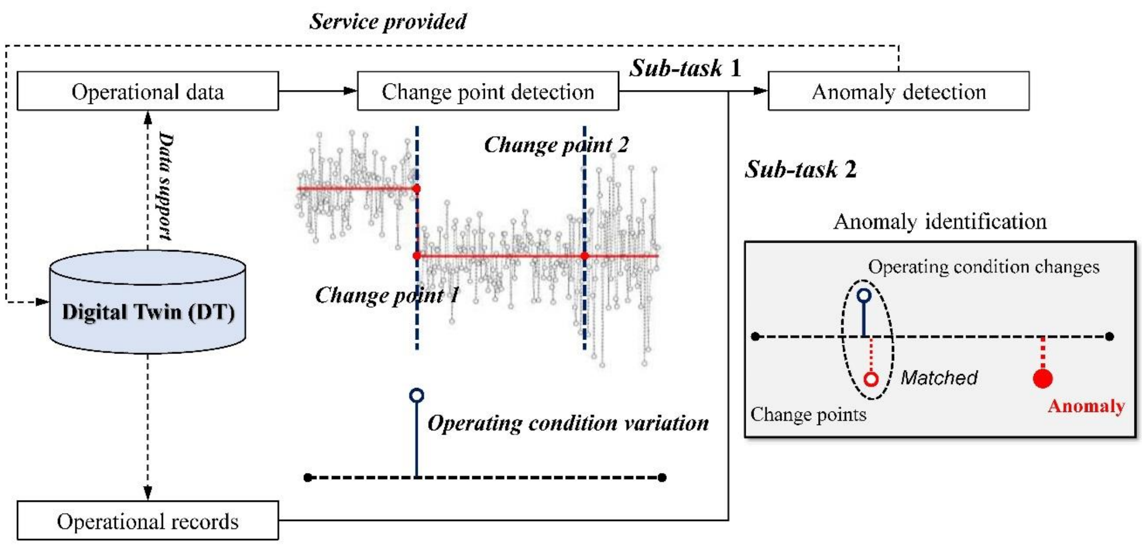

Figure 8 Procedure of anomaly detection for asset monitoring

BOCPD approach is adopted in this procedure because it does not require any prior knowledge of pre-change or post-change operation processes, which is exactly the case of anomaly detection for building assets. With BOCPD algorithm, the objective is, given a sequence of operational data $\mathbf{x}=\{x_1,\cdots,x_t,\cdots\}$ collected from a specific asset, to compute the posterior probability distribution $p(r_t \mid \mathbf{x})$ over the run length r_t, referring to the number of observations since the last found change point. The run length is truncated to 0 if a change point is identified, otherwise, the run length increases by one as the observation of new data points x_t comes. It implies that the last change point occurs at the time $t-r_t$ and the set of observed data associated with the current run is $\mathbf{x}_t^r=\{x_{t-r_t+1},\cdots,x_t\}$. Under the Bayesian framework, the posterior distribution r_t can be expanded using Bayes law:

$$
\begin{aligned}
p(r_t \mid x_{1:t}) \propto p(r_t, x_{1:t}) &= \sum_{r_{t-1}} p(r_t, r_{t-1}, x_{1:t}) \\
&= \sum_{r_{t-1}} p(r_t, x_t \mid r_{t-1}, x_{1:t-1})\, p(r_{t-1}, x_{1:t-1}) \\
&= \sum_{r_{t-1}} p(r_t \mid r_{t-1})\, p(x_t \mid \mathbf{x}_t^r)\, p(r_{t-1}, x_{1:t-1})
\end{aligned}
\tag{1}
$$

Note that $p(r_t \mid x_{1:t})$ becomes the function of $p(r_{t-1} \mid x_{1:t-1})$, which mean that the distribution of run length can be calculated in a recursive fashion, suitable for the online update using a recursive message-passing scheme. The scheme updates the posterior over the run length based on two calculations, the change point prior $p(r_t \mid r_{t-1})$ and the predictive distribution $p(x_t \mid \mathbf{x}_t^r)$ over the new observation given the most recent data points in the single run, respectively.

For simplicity, the assumption is made that the length of each run follows an exponential distribution and the prior probability of a change point is given by the pre-specified hazard rate h independent of r_t, and $p\left(r_t \mid r_{t-1}\right)=h$ if the run length resets while $p\left(r_t \mid r_{t-1}\right)=1-h$ when $r_t = r_{t+1}+1$. The predictive distribution $p\left(x_t \mid \mathbf{x}_t^r\right)$ depends only on the knowledge of the generative process $\mathbf{x}_t^r$ that was active before the last identified change point. Specifically, the predictive distributions $p\left(x_t \mid \mathbf{x}_t^r\right)$ can be conveniently described by a finite number of sufficient statistics if generative distributions are members of the conjugate-exponential family likelihoods. Assuming that the generative process $\mathbf{x}_t^r$ follows a Gaussian distribution with unknown mean θ and variance λ. In this case, a joint conjugate prior on θ and λ can be expressed in a general form of Normal-Gamma distribution with the prior hyper-parameter set $\eta_0 = \left\{\mu_0, \kappa_0, \alpha_0, \beta_0\right\}$:

$$p\left(\theta, \lambda \mid \mu_0, \kappa_0, \alpha_0, \beta_0\right) = \frac{\beta_0^{\alpha_0}\sqrt{\kappa_0}}{\Gamma\left(\alpha_0\right)\sqrt{2\pi}} \lambda^{\alpha_0 - \frac{1}{2}} \exp\left(-\frac{\lambda}{2}\left[\kappa_0\left(\theta-\mu_0\right)^2 + 2\beta_0\right]\right) \tag{2}$$

As new observations arrive incrementally, the hyper-parameter set updates in the form as follows:

$$\begin{aligned}
\mu_t &= \frac{\kappa_0\mu_0 + t\overline{\mathbf{x}}_t^r}{\kappa_0 + t} \\
\kappa_t &= \kappa_0 + t \\
\alpha_t &= \alpha_0 + \frac{t}{2} \\
\beta_t &= \beta_0 + \frac{1}{2}\sum_{i=t-r_t+1}^{t}\left(x_i - \overline{\mathbf{x}}_t^r\right)^2 + \frac{\kappa_0 t\left(\overline{\mathbf{x}}_t^r - \mu_0\right)^2}{2\left(\kappa_0 + t\right)}
\end{aligned} \tag{3}$$

Following the inference, the posterior predictive distribution $p\left(x_t \mid \mathbf{x}_t^r\right)$ follows a generalized student's t-distribution with mean μ_t, variance $\beta_t\left(\kappa_t+1\right)/\alpha_t\kappa_t$ and $2\alpha_t$ degree of freedom.

5. Case Study

5.1 DT construction and data integration

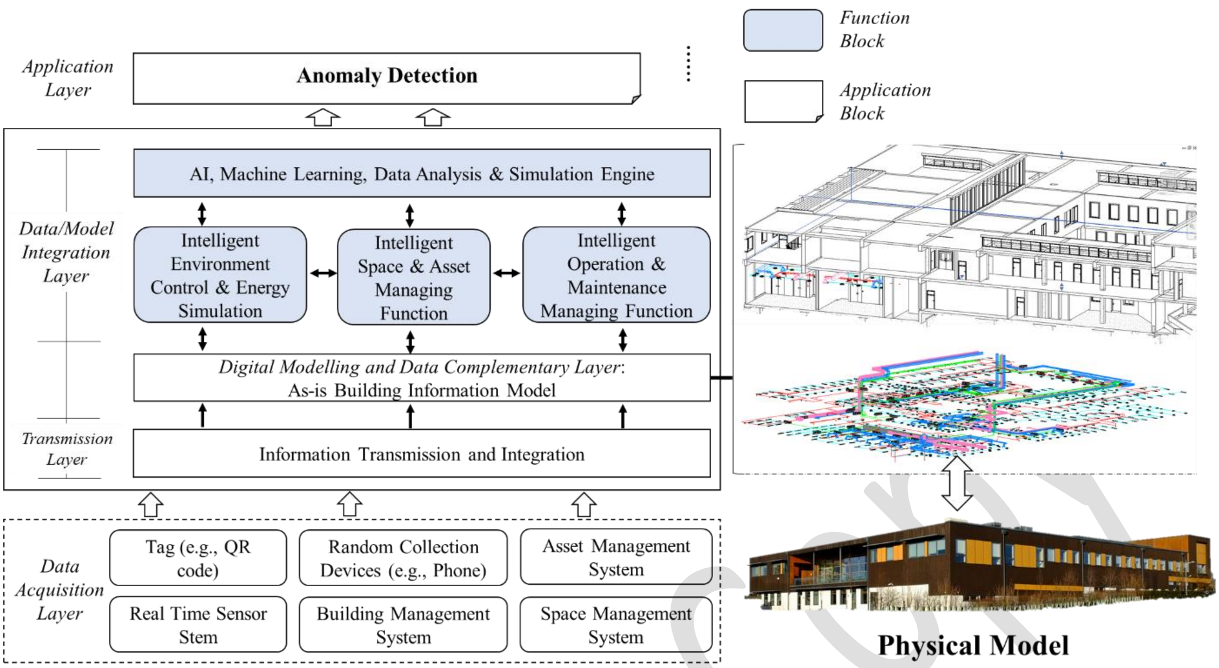

Figure 9 The developed building DT (modified from [44] and [64])

The pilot evaluation study of the proposed building DT was conducted in the Institute for Manufacturing (IfM) building at the West Cambridge site of the University of Cambridge. The IfM building is a 3-storey building and stands over a 40000-square-foot comprehensive area, including teaching, study, office, research and laboratory spaces. Based on the designed architecture [9,64], the developed IfM building DT includes five layers, integrates various data resources and also supports anomaly detection (Fig.9). The objective of this case study is to demonstrate how the designed data structure can contribute to the data integration of a dynamic DT of existing buildings, to support its anomaly detection function and further to explore the opportunities and challenges.

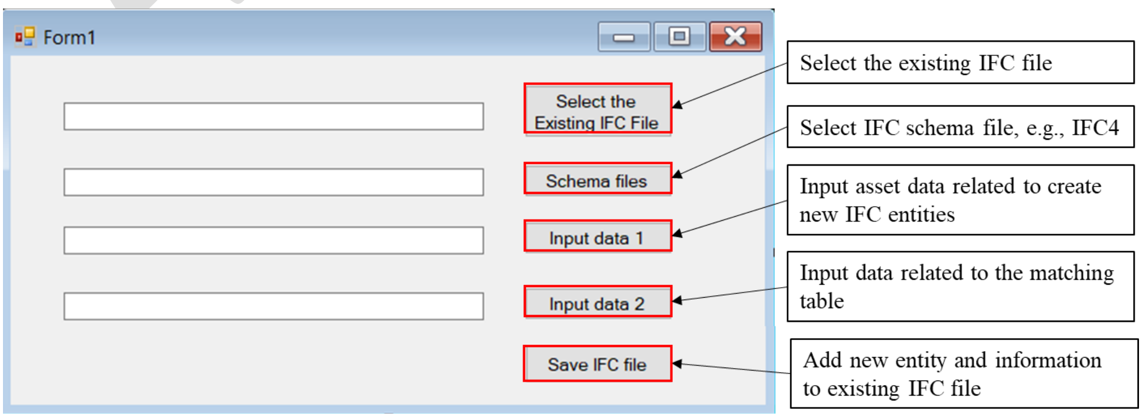

Figure 10 Application development

Firstly, the IFC extension application (as shown in Fig.10) is developed for creating new IFC files in accordance with the existing ones. Three functions are included in this application:

1). Add missing components (e.g., *IfcPump*) in the existing IFC file;

2). Save needed information of matching tables as a reference/backup in the existing IFC file;

3). Add and create additional entities in the existing IFC file.

Based on the updated IFC file, an *IfcObject* matching table used for data integration is created to describe the interconnection between the BIM object Globally Unique Identifier (GUID) and corresponding item ID from different data sources (e.g., BMS and sensor system). As shown in Figure 11, when a data item (saved in distributed BMS or sensor system) needs to be integrated or queried for anomaly detection in the upper layer, the *IfcObject* matching table provides linking bridges between the targeted BIM object (GUID) and the corresponding item ID in BMS, and similarly between the BIM object (GUID) and the required sensor ID in the sensor system. Through the matching process, the matched item ID is used as a primary key (PK) in the designed data schema for searching the required data. Through the GUID in the *IfcObject* matching table and querying matched item ID number, the required data would be searched automatically by their unique item ID as primary key and further refined using sort key (SK). Similarly, required sensor data would also be queried. In this way, data needed for anomaly detection would be queried automatically and linked to their corresponding BIM object. This enables IFC and other data sources (e.g., BMS) to be saved separately in a distributed approach. To keep the consistency of the data, only the *IfcObject/IfcSpace* matching table needs to be maintained, which achieves effective CRUD (Create, Retrieve, Update, Delete).

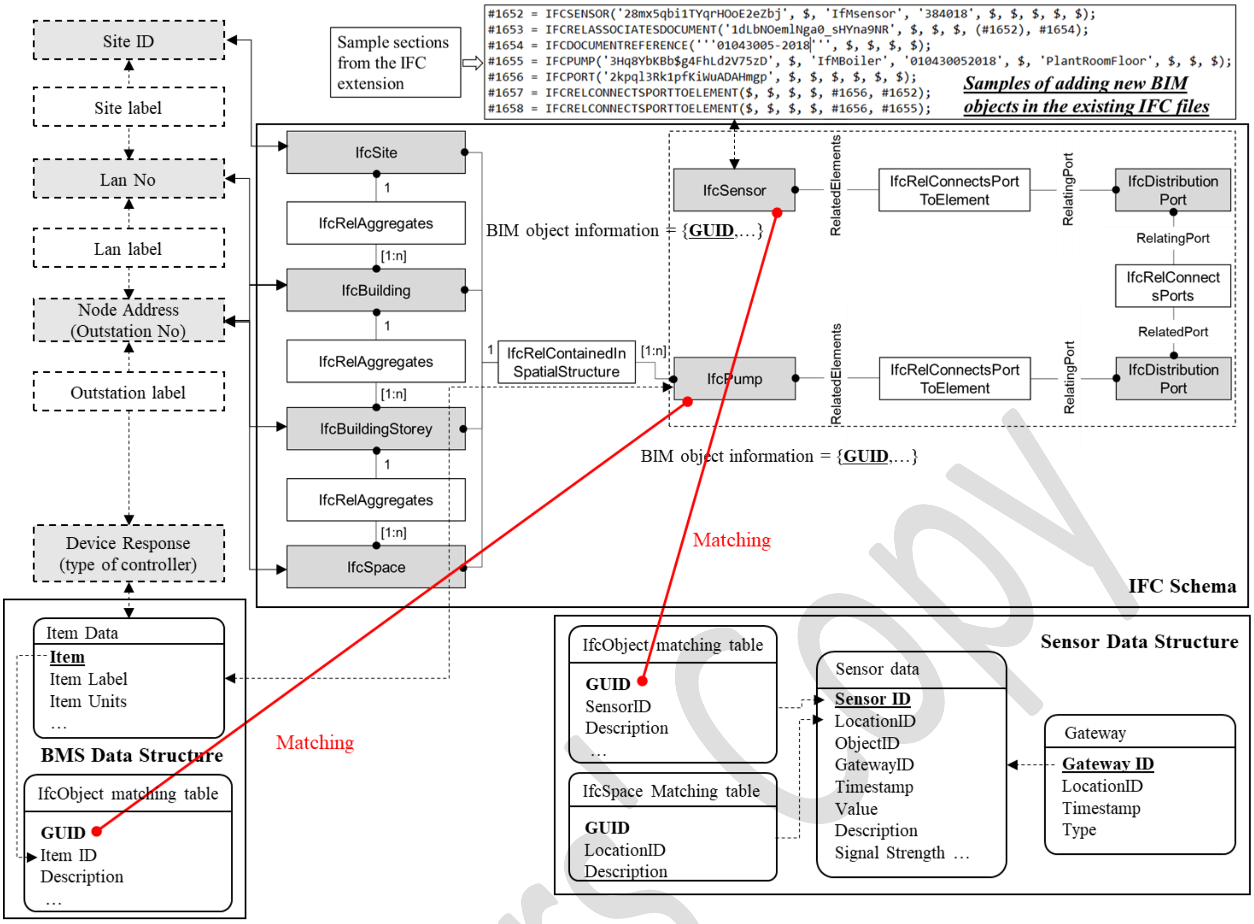

Figure 11 The IFC schema mapping with other data resources (BMS and sensor system)

5.2 Anomaly detection and comparative analysis

In this section, the application of the proposed anomaly detection procedure is illustrated on monitoring of two centrifugal pumps, and the experiment results are presented. Two pumps of the same specifications are installed in the plant room of the 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 failures like defective bearing, sealing, or defect on impeller and cavitation could result in negative and even catastrophic consequences, such as abnormal noises, rotating unbalance, shaft breakage. The most revealing and widely accepted diagnostic information on the mechanical condition of the centrifugal pump is the vibration measurements, because vibration data contains abundant information about machinery running states with reasonable sensing costs [70]. Because the vibrations are transferred from the 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. Besides the vibration data, data from

BMS, such as pump on-duty flag bit, is integrated as external asset operation information for filtering the identified contextual anomalies.

With the help of embedded sensor systems, a long period of averaged vibration frequency data is integrated into the DT demonstrator, which makes it possible to continuously conduct a tentative diagnosis for the pumps' health condition. The data include the response to both scheduled operating shutdown and a real anomaly causing strong abnormal noises. Two datasets with a sampling time of one hour are picked to examine and compare the relative performance of the conventional cumulative sum control charts (CUSUM) with the proposed method. In the first case, the studied centrifugal pump 1 undergoes a scheduled shutdown due to the UK bank holiday. The period of data starts from the 5th December of 2018 and lasts until 1st January of 2019 (4 weeks). Fig.12 shows the recorded vibration frequency time series within a 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.

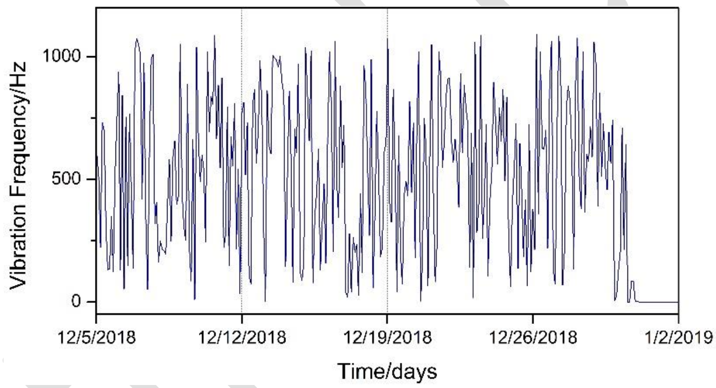

Figure 12 Vibration frequency sequence in the pump shutdown scenario

The intuitive derivation of two-sided CUSUM algorithm is first utilized to detect the shutdown induced change point in the recorded data. The detection result is illustrated by Fig.13. The blue upper-sided CUSUM chart detects the increase in the featured vibration frequency, while the red lower-sided CUSUM chart detects the decrease in the frequency. As shown in the Fig.13, the CUSUM based detector successfully locates the frequency change point corresponding to the shutdown scenario within a reasonable time. However, two false alarms events are generated in this period.

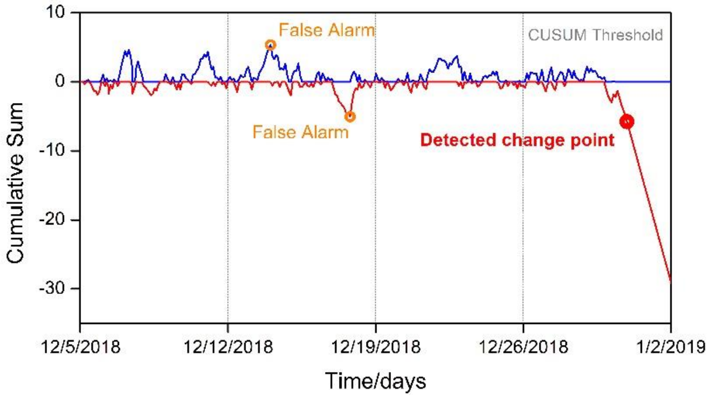

Figure 13 Detection of the shutdown event by CUSUM procedure

Then, the proposed BOCPD based procedure is adopted to detect the change points for the same data sequence. Fig.14 depicts the output of the BOCPD based procedure when applied to the pump shutdown event. The top plot labels the change point detection result, in which the vertical dashed blue line represents the identified shutdown time using BOCPD, and the black cross marks the point detected by CUSUM. The detected change point times using CUSUM and BOCPD are almost identical. But BOCPD based method effectively avoids the raised false alarm. The red solid line reveals the local maximum a posterior run length estimation result, while the blue dashed line marks the most probable run length considering the continuity of the run length. Although the local optimal run length shows some spikes, the BOCPD is able to compensate for the side effects caused by occasional measurement errors. This is the key point to the reduction of the false alarm rate.

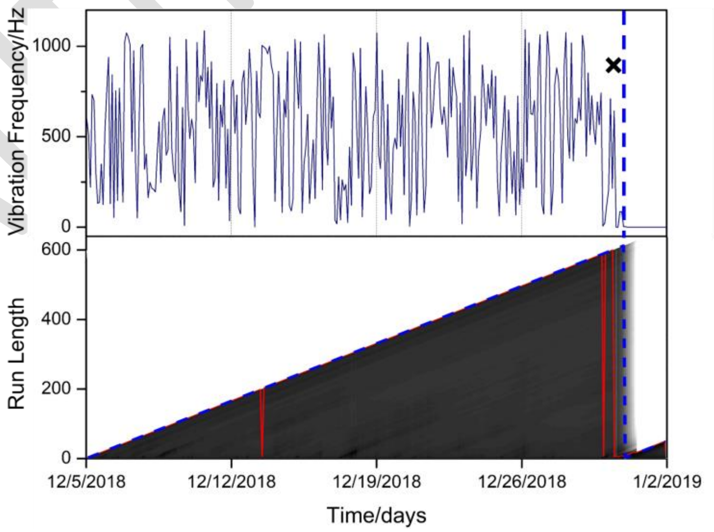

Figure 14 Detection of the shutdown event by BOCPD based procedure

For the second case, one of the two pumps undergoes a highly suspicious anomaly causing a strong abnormal degree of noise, while the other one works properly. Here, an artificial dataset is generated by combining 14 days vibration frequency data from the normal pump with 14 days data from the anomalous pump (from 9th July to 22nd July in 2018). Fig.15 shows the generated vibration frequency time series within a given period. Different from the shutdown scenario, it is hard to distinguish the difference between the vibration of normal and anomalous pumps by unaided eyes. Therefore, both CUSUM and BOCPD are utilized to detect the change point between two kinds of vibration frequencies.

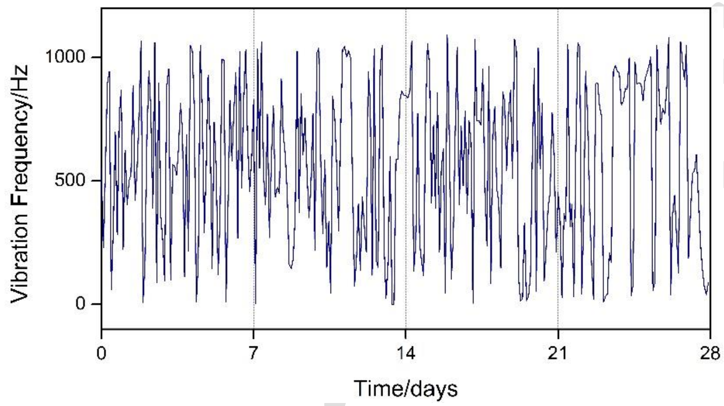

Figure 15 Vibration frequency sequence in the pump anomalous scenario

The detection result using the CUSUM control chart is illustrated by Fig.16. The procedure successfully detects the vibration frequency deviation with a considerable delay of almost a week. It is because the vibration frequency is not informative enough, thus it only offers a very rough diagnosis for the working condition of the pump. A longer time is needed to accumulate the anomaly indicative frequency deviations before reaching the determined threshold defined in the CUSUM chart.

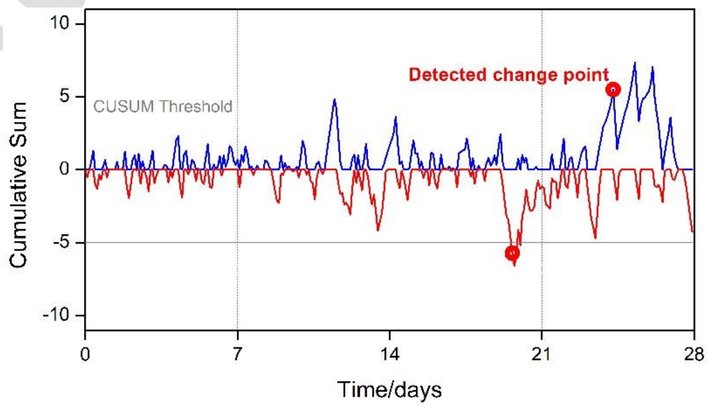

Figure 16 Detection of the pump anomalous event by CUSUM procedure

Similarly, the BOCPD based procedure is utilized for the same data sequence. Fig.17 depicts the output of the BOCPD based approach when applied to the pump anomalous event. Obviously, the BOCPD procedure shows a better capability of detecting changes with a little time delay when compared to CUSUM. However, as shown in the bottom plot, the red cross labels the awareness time. The advantage of BOCPD based procedure is that although there is a slight delay before the anomaly of pumps are recognized, actual change point time can be uniquely pin pointed when subsequent indicative data is available. For the cross-over match process, a more precise change point contributes to the matching between symptoms and corresponding normal operations.

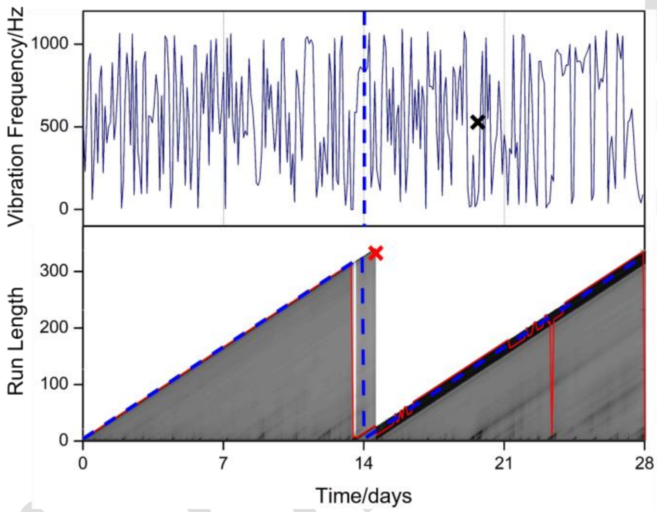

Figure 17 Detection of the pump anomalous event by BOCPD procedure

5.3 DT Platform Design and Visualization

On the basis of the anomaly detection capability established in section 5.2 and data integration in section 5.1, the DT platform provides the asset monitoring service to facility managers and other related stakeholders by interpreting professional knowledge embedded in the established anomaly detection module and practically enabling interaction between the physical and digital world. Although the DT properly manages and integrates multi-source data through IFC schema and intelligently analyses these data in a systematic way, the ultimate objective of the DT platform is to provide intuitional information visualization and decision support to FM professionals. In order to establish the DT platform, Autodesk Revit was used to develop the RVT model and then export it to IFC files. The platform was developed based on AWS DynamoDB, Autodesk forge API and web-based program design (i.e., .Net) using C# and Java script [9,64]. Taking advantage of these tools, the asset monitoring service is enabled in the developed DT platform (as shown in Fig.18). With the capability to store and analyse BIM object related data collected by heterogeneous data sources, the embedded DT instance implements the intelligent extraction of pump relevant data and triggers the alarm once the anomaly detection procedure finds any possible anomalous behaviour for the studied pump.

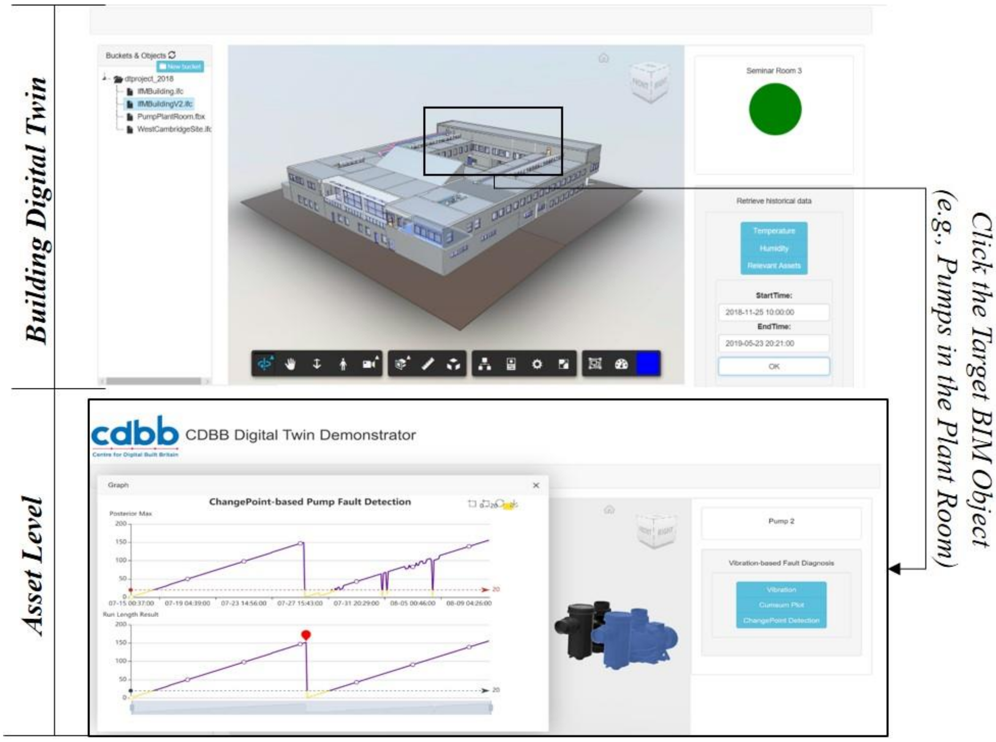

Figure 18 Asset Monitoring service provided by DT platform

6. Discussion

In order to reveal the anomalous behaviour of assets in a timely manner, and take preventative actions before severe and even catastrophic consequences happen, an anomaly detection system for asset monitoring during the O&M phase is urgently needed. In spite of great efforts devoted to fulfil anomaly detection automatically, the anomaly detection task of building assets is mainly completed manually by experienced FM professionals. Advanced analytical tools, including those based on machine learning or artificial intelligence, should be capable of distinguishing between different patterns behind the operational data. However, the real challenge is that single source data couldn't provide a holistic view under the continuously changing working condition of typical assets. In this study, an anomaly detection procedure for circulating pumps is discussed. Typically, vibration sensors are mounted on the pumps to monitor the vibration frequency, which indicates their working condition. It is easy to identify that the characteristic of the pump vibration gradually drifts with the changes of working loads/conditions. For instance, the vibration characteristic during peak loading hours is different from that during valley loading hours. However, neither of these two characteristics manifest the anomalous behaviour of pumps. That is to say, classical point anomaly detection does contribute to clarifying the asset behavioural changes, but still lacks enough explanatory factors that distinguish anomalous behaviours from normal ones. To solve this, one of the possible strategies is to train an unsupervised or one-class classifier using a refined normal dataset under various loading scenarios [72]. Additional data and information, such as the BMS data, is necessary to divide the historical data into normal and anomalous parts. However, to make the classifier generalized enough, massive data under a large number of normal working conditions is required for training, which is impractical. Given all the practical constrains, another strategy adopted here is to temporally identify change point raised non-stationary events, which manifest as variations in the generative parameters of the data sequence. Subsequently, BMS in this case, needs to be integrated to eliminate the change points raised by normal operations and leave anomaly raised change points as the trigger for following-up early warning. Specifically, the matching between logged operating condition variations and detected change point determines those eliminated change points. The matching can be simple or complex, depending on the accuracy of the change point detection algorithm in pin-pointing the time of change points or non-stationary events. It is verified in the case study that the Bayesian on-line change point detection algorithm is capable of accurately recognizing the time of change, even though the awareness time would be slightly delayed. It makes simple cross-over matching sufficient for the pump anomaly detection module.

It is worth noting that the capability to store, manipulate, exchange and analyse BIM objects (pumps in this case) related data collected by heterogeneous data sources is the core competence of the DT-enabled anomaly detection system of asset monitoring. In particular, DT improves data management efficiency, and makes it easier to integrate data from autonomous, disparate and heterogeneous sources. Traditionally, the efficient execution of queries to extract the data from disparate systems is non-trivial. With the help of the standardized IFC schema, an object-oriented and semantic BIM representation is presented that includes components, attributes, properties, relationships, and most importantly linkages with multiple data resources. In this way, exchanging information across data source boundaries is enabled using IFC schema in the DT platform.

Although the proposed anomaly detection procedure can realize asset monitoring, as verified in the case study, we must realize that considering the budget constraints, it is impossible to monitor every single asset within such a complicated building system at a fine granularity.

Only critical assets, for instance, the pumps in the case study, have corresponding monitoring data in either sensor system or BMS. For those noncritical assets, such as valves or pipelines, no relevant data is explicitly linked to the specific object. However, the condition of these noncritical assets can be monitored through the quality of service (QoS)/performance provided by building systems. For instance, the room temperature would drop significantly in winter if the radiator valve fails to open properly. Therefore, in addition to the anomaly detection system of asset monitoring, indoor environment monitoring system also needs to be developed under the framework of DT to enable better understanding of the working conditions of various building assets.

7. Conclusions

In order to provide a comprehensive asset monitoring solution in the building O&M phase, a DT-enabled anomaly detection system was developed in this study. The developed system is useful for detecting anomalies of building assets and can be crucial for daily O&M management. It not only demonstrates the application of the designed IFC extension and BOCPD in detecting suspicious anomalies of pumps, but also contributes to research advancement by:

- Proposing a new DT-based anomaly detection process flow, realizing effective data integration and information search, facilitating decision making and automating the anomaly detection process;
- Designing the structure of data integration based on IFC extension in O&M management for heterogeneous operational data storage, exchange, query and update;
- Identifying the capability of distinguishing asset behavioural changes caused by normal operating condition variations or true anomalies using conventional anomaly detection;
- Adopting a Bayesian change point detection methodology that handles the contextual features of behavioural data to identify and filter asset anomalies through cross-referencing with external operation information.

A case study using the pumps in HVAC system was used to evaluate and demonstrate the effectiveness of the proposed framework. The results indicated that the provided solution realized a continuous condition monitoring of building assets (e.g., pumps) and also contributed to efficient and automated asset monitoring in the daily O&M management.

This research contributes to the body of knowledge by developing a novel system for future researchers to systematically and intelligently monitor assets based on DTs. In future work, we will keep working on information integration strategies (e.g., expert experience) through working with Estate Management department in this University, extend building assets to broader city assets and investigate more practical applications of the DTs development in supporting the wider management activities and services.

Acknowledgement

This research that contributed to this paper was supported by the Centre for Digital Built Britain (CDBB) with funding provided through the Government's modern industrial strategy by Innovate UK, part of UK Research & Innovation. It was also partly funded by the EPSRC/Innovate UK Centre for Smart Infrastructure and Construction (Grant Number EP/N021614/1).

Reference

[1] NRC, Stewardship of federal facilities, A Proactive Strategy for Managing the nation's Public Assets, National Research Council, National Academies Press, Washington, DC, 1998.

[2] E.M. Wetzel, W.Y. Thabet, The use of a BIM-based framework to support safe facility management processes, Automation in Construction 60 (2015) 12-24, http://doi.org/10.1016/j.autcon.2015.09.004.

[3] Q. Lu, L. Chen, S. Lee, X. Zhao, Activity theory-based analysis of BIM implementation in building O&M and first response, Automation in Construction 85 (2018) 317-332, https://doi.org/10.1016/j.autcon.2017.10.017.

[4] D. Sapp, Whole building design guide, (2015), Last accessed October 1, 2016, from http://www.wbdg.org/om/om.php.

[5] L. Ding, R. Drogemuller, P. Akhurst, R. Hough, S. Bull, C. Linning, Towards sustainable facilities management, Peter Newton, Keith Hampson, Robin Drogemuller (Eds.), Technology, Design and Process Innovation in the Built Environment, Taylor & Francis, Oxon, Abingdon (2009), pp. 373–392, http://eprints.qut.edu.au/20926/.

[6] I. Motawa, A. Almarshad, A knowledge-based BIM system for building maintenance, Automation in construction 29 (2013) 173-182, http://doi.org/10.1016/j.autcon.2012.09.008.

[7] P. Parsanezhad, J. Dimyadi, Effective facility management and operations via a BIM based integrated information system, CIB Facilities Management (CFM) 2014 Conference, Copenhagen, Denmark, 2014, pp. 8, Last accessed January 10, 2018 from http://www.cfm.dtu.dk/english/CIB-Conference.

[8] Z. Shi, W. O'Brien, Development and implementation of automated fault detection and diagnostics for building systems: A review, Automation in Construction 104(2019) 215-229, https://doi.org/10.1016/j.autcon.2019.04.002.

[9] Q. Lu, A.K. Parlikad, P. Woodall, G.D. Ranasinghe, J. Heaton, Developing a dynamic digital twin at a building level: using Cambridge campus as case study, International Conference on Smart Infrastructure and Construction (ICSIC), Cambridge, UK, 2019.

[10] Gartner, Prepare for the Impact of Digital Twins, (2017), Last accessed April 25, 2019, from https://www.gartner.com/smarterwithgartner/prepare-for-the-impact-of-digital-twins/.

[11] GE Digital, Digital Twins: The Bridge Between Industrial Assets and the Digital World, (2017), Last accessed April 25, 2019, from https://www.ge.com/digital/blog/digital-twins-bridge-between-industrial-assets-and-digital-world.

[12] National Infrastructure Commission (NIC), Data for the public good, (2017), Last accessed April 25, 2019, from https://www.nic.org.uk/wp-content/uploads/Data-for-the-Public-Good-NIC-Report.pdf.

[13] A. Costin, A. Shaak, J. Teizer, Development of a navigational algorithm in BIM for effective utility maintenance management of facilities equipped with passive RFID, ASCE Computing in Civil Engineering, Los Angeles, CA, 2013, pp. 653–660, http://dx.doi.org/10.1061/9780784413029.082.

[14] W. Shen, Q. Hao, Y. Xue, A loosely coupled system integration approach for decision support in facility management and maintenance, Automation in construction 25 (2012) 41-48, https://doi.org/10.1016/j.autcon.2012.04.003.

[15] M. Dibley, H. Li, Y. Rezgui, J. Miles, An ontology framework for intelligent sensor-based building monitoring, Automation in Construction 28 (2012) 1-14, https://doi.org/10.1016/j.autcon.2012.05.018.

[16] J. Lee, Y. Jeong, Y.S. Oh, J.C. Lee, N. Ahn, J. Lee, S.H. Yoon, An integrated approach to intelligent urban facilities management for real-time emergency response, Automation in construction 30 (2013) 256-264, https://doi.org/10.1016/j.autcon.2012.11.008.

[17] H.L. Chi, S.C. Kang, X. Wang, Research trends and opportunities of augmented reality applications in architecture, engineering, and construction, Automation in construction 33 (2013) 116-122, https://doi.org/10.1016/j.autcon.2012.12.017.

[18] B.R. Kyle, D.J. Vanier, B. Kosovac, T.M. Froese, Z. Lounis, Visualizer: an interactive, graphical, decision-support tool for service life prediction for asset managers, Proceeding of 9th International Conference on Durability of Building Materials and Components, Brisbance,

2002, pp. 17–20, Last accessed January 01, 2016 from www.irbnet.de/daten/iconda/CIB9286.pdf.

[19] B. Succar, Building information modelling framework: A research and delivery foundation for industry stakeholders, Automation in Construction 18(3) (2009) 357-375, https://doi.org/10.1016/j.autcon.2008.10.003.

[20] W. Chen, K. Chen, J.C. Cheng, Q. Wang, V.J. Gan, BIM-based framework for automatic scheduling of facility maintenance work orders, Automation in construction 91 (2018) 15-30, https://doi.org/10.1016/j.autcon.2018.03.007.

[21] E.M. Wetzel, W.Y. Thabet, A case study towards transferring relevant safety information for facilities maintenance using BIM, Journal of information technology in construction (ITcon) 23(3) (2018) 53-74, ISSN 1874-4753.

[22] N.D. Aziz, A.H. Nawawi, N.R.M. Ariff, ICT evolution in facilities management (FM): building information modelling (BIM) as the latest technology, Procedia-social and behavioral sciences 234 (2016) 363-371, https://doi.org/10.1016/j.sbspro.2016.10.253.

[23] IBM Corporation, Implementation Guide for Integrated Workplace Management Software, IBM Corporation, US, 2013, Last accessed January 10, 2018 from http://www-01.ibm.com/common/ssi/cgi-bin/ssialias?subtype=WH&infotype=SA&appname=SWGE_TI_EA_USEN&htmlfid=TIW14165USEN&attachment=TIW14165USEN.PDF.

[24] S. Lin, J. Gao, A. Koronios, Key data quality issues for enterprise asset management in engineering organisations, Enterprise Asset Management in Engineering Organisations (IJEBM) 4 (1) (2006) 96–110.

[25] P. Teicholz, BIM for Facility Managers, John Wiley & Sons, New Jersey, 2013, ISBN-13: 978-1118382813.

[26] V. Aspurez, P. Lewis, Case study 3: USC school of cinematic arts, BIM for Facility Managers, Wiley, Hoboken, NJ, 2013, pp.185-232.

[27] Y.C. Lin, Y.C. Su, Developing mobile-and BIM-based integrated visual facility maintenance management system, The scientific world journal (2013), http://dx.doi.org/10.1155/2013/124249.

[28] Y.C. Lin, Y.C. Su, Y.P. Chen, Developing mobile BIM/2D barcode-based automated facility management system, The Scientific World Journal (2014), http://dx.doi.org/10.1155/2014/374735.

[29] M. Arslan, Z. Riaz, S. Munawar, Building Information Modeling (BIM) Enabled Facilities Management Using Hadoop Architecture., Portland International Conference, Management of

Engineering and Technology (PICMET), IEEE, Portland, 2017, pp.1-7, 10.23919/PICMET.2017.8125462.

[30] K. Suprabhas, H.N. Dib, Integration of BIM and utility sensor data for facilities management, ASCE International Workshop on Computing in Civil Engineering 2017, Seattle, Washington, USA, 2017, pp. 26-33, https://doi.org/10.1061/9780784480823.004.

[31] H. Schevers, J. Mitchell, P. Akhurst, D. Marchant, S. Bull, K. McDonald, R. Drogemuller, Towards digital facility modelling for sydney opera house using IFC and semantic web technology, Journal of information technology in construction (ITcon) 12 (2007) 347–362, ISSN: 1874-4753.

[32] E.A. Pärn, D.J. Edwards, M.C.P. Sing, The building information modelling trajectory in facilities management: A review, Automation in construction 75 (2017) 45-55, https://doi.org/10.1016/j.autcon.2016.12.003.

[33] B. Becerik-Gerber, F. Jazizadeh, N. Li, G. Calis, Application areas and data requirements for BIM-enabled facilities management, Journal of construction engineering and management 138(3) (2011) 431-442, https://doi.org/10.1061/(ASCE)CO.1943-7862.0000433.

[34] J. Patacas, N. Dawood, V. Vukovic, M. Kassem, BIM for facilities management: evaluating BIM standards in asset register creation and service life planning, Journal of information technology in construction (ITcon) 20(10) (2015) 313-318, ISSN: 1874-4753.

[35] T.W. Kang, H.S. Choi, BIM perspective definition metadata for interworking facility management data, Advanced engineering informatics 29(4) (2015) 958-970, https://doi.org/10.1016/j.aei.2015.09.004/.

[36] H.B. Cavka, S. Staub-French, E.A. Poirier, Developing owner information requirements for BIM-enabled project delivery and asset management, Automation in construction 83 (2017) 169-183, https://doi.org/10.1016/j.autcon.2017.08.006.

[37] C. Nicolle, C. Cruz, Semantic building information model and multimedia for facility management, International Conference on Web Information Systems and Technologies, Springer, Berlin, Heidelberg, 2010, pp. 14-29, https://doi.org/10.1007/978-3-642-22810-0_2.

[38] J. Korpela, R. Miettinen, T. Salmikivi, J. Ihalainen, The challenges and potentials of utilizing building information modelling in facility management: the case of the Center for Properties and Facilities of the University of Helsinki, Construction management and economics 33(1) (2015) 3-17, https://doi.org/10.1080/01446193.2015.1016540.

[39] S.O. Alvarez-Romero, Use of Building Information Modeling Technology in the Integration of the Handover Process and Facilities Management, Worcester Polytechnic Institute, 2014, Dissertation, Last accessed March 10, 2018 from https://www.wpi.edu/

Pubs/ETD/Available/etd-090914…/Disertation_final_SA.pdf.

[40] H.M. Chen, C.C. Hou, Y.H. Wang, A 3D visualized expert system for maintenance and management of existing building facilities using reliability-based method, Expert Systems with Applications 40(1) (2013) 287-299, https://doi.org/10.1016/j.eswa.2012.07.045.

[41] P.E. Love, J. Matthews, I. Simpson, A. Hill, O.A. Olatunji, A benefits realization management building information modeling framework for asset owners, Automation in construction 37 (2014) 1-10, https://doi.org/10.1016/j.autcon.2013.09.007.

[42] The State of Wisconsin, Digital facility management information handover, Current DSF Practices Industry-wide Movement Future Directions, a Research, Findings and Recommendations Report, Vol. Jul 15, 2011, Last accessed December 01, 2017 from ftp://doaftp1380.wi.gov/master_spec/Digital%20FM%20Handover/FM%20Findings&RecRpt.pdf.

[43] UNITEC's Integrated Information System, BIM As An Information Sharing Resource For Facilities Management And Operations, UNITEC, Last accessed January 01, 2018 from https:// www.building.govt.nz/assets/Uploads/projects-and-consents/building-information-modelling/nz-bim-case-study-5-unitec.pdf.

[44] I.F. Cruz, H. Xiao, Ontology Driven Data Integration in Heterogeneous Networks, Complex Systems in Knowledge-based Environments: Theory, Models and Applications, Springer, Heidelberg, 2009, 75–98, https://doi.org/10.1007/978-3-540-88075-2_4.

[45] A. Hassanain, T. Froese, D. Vanier, Implementation of a distributed, model-based integrated asset management system, Journal of Information Technology in Construction (ITcon) 8(10) (2003) 119–134, ISSN: 1874-4753.

[46] V. Chandola, A. Banerjee, V. Kumar, Anomaly detection: A survey, ACM Computing Surveys (CSUR) 41(3) (2009) 15, https://doi.org/10.1145/1541880.1541882.

[47] A. Capozzoli, F. Lauro, I. Khan. Fault detection analysis using data mining techniques for a cluster of smart office buildings, Expert Systems with Applications 42(9) 4324-4338, https://doi.org/10.1016/j.eswa.2015.01.010.

[48] J.E. Seem, Using intelligent data analysis to detect abnormal energy consumption in buildings. Energy and buildings, 39(1) 52-58, https://doi.org/10.1016/j.enbuild.2006.03.033.

[49] X. Li, C.P. Bowers, T. Schnier, Classification of energy consumption in buildings with outlier detection, IEEE Transactions on Industrial Electronics 57(11) 3639-3644, https://doi.org/10.1109/TIE.2009.2027926.

[50] M. Molina-Solana, M. Ros, M.D. Ruiz, J. Gomez-Romero, M.J. Martin-Bautista, Data science for building energy management: A review, Renewable and Sustainable Energy Reviews 70(2017) 598-609, https://doi.org/10.1016/j.rser.2016.11.132.

[51] D. Jacob, S. Dietz, S. Komhard, C. Neumann, S. Herkel, Black-box models for fault detection and performance monitoring of buildings, Journal of Building Performance Simulation 3(1) 53-62, https://doi.org/10.1080/19401490903414454.

[52] F. Xiao, C. Fan, Data mining in building automation system for improving building operational performance, Energy and buildings 75(2014) 109-118, https://doi.org/10.1016/j.enbuild.2014.02.005.

[53] C. Fan, F. Xiao, C. Yan, A framework for knowledge discovery in massive building automation data and its application in building diagnostics, Automation in Construction 50(2015) 81-90, https://doi.org/10.1016/j.autcon.2014.12.006.

[54] Z.J. Yu, F. Haghighat, B.C. Fung, L. Zhou, A novel methodology for knowledge discovery through mining associations between building operational data, Energy and Buildings 47(2012) 430-440, https://doi.org/10.1016/j.enbuild.2011.12.018.

[55] D.F.M. Cabrera, H. Zareipour, Data association mining for identifying lighting energy waste patterns in educational institutes, Energy and Buildings 62(2013) 210-216, https://doi.org/10.1016/j.enbuild.2013.02.049.

[56] F. Xiao, C. Fan, Data mining in building automation system for improving building operational performance, Energy and Buildings 75(11) 109-118, https://doi.org/10.1016/j.enbuild.2014.02.005.

[57] C. Fan, F. Xiao, Z. Li, J. Wang, Unsupervised data analytics in mining big building operational data for energy efficiency enhancement: A review, Energy and Buildings 15(2018) 296-308, https://doi.org/10.1016/j.enbuild.2017.11.008.

[58] A.G. Tartakovsky, A.S. Polunchenko, G. Sokolov, Efficient Computer Network Anomaly Detection by Changepoint Detection Methods, IEEE Journal of Selected Topics in Signal Processing 7(1) 4-11, https://doi.org/10.1109/JSTSP.2012.2233713.

[59] S. Touzani, V. Ravache, E. Crowe, J. Granderson, Statistical change detection of building energy consumption: Applications to savings estimation. Energy and Buildings 185(2019) 123-136, https://doi.org/10.1016/j.enbuild.2018.12.020.

[60] H.B. Gunay, W. Shen, G. Newsham, Data analytics to improve building performance: A critical review, Automation in Construction, 97(2019) 96-109, https://doi.org/10.1016/j.autcon.2018.10.020.

[61] T.W. Kang, C.H. Hong, A study on software architecture for effective BIM/GIS-based

facility management data integration, Automation in Construction, 54 (2015) 25–38, http://dx. doi.org/10.1016/j.autcon.2015.03.019.

[62] A. Costa, M.M. Keane, J.I. Torrens, E. Corry, Building operation and energy performance: Monitoring, analysis and optimisation toolkit, Applied Energy, 101(2013) 310-316, https://doi.org/10.1016/j.apenergy.2011.10.037.

[63] A. Motamedi, A. Hammad, Y. Asen, Knowledge-assisted BIM-based visual analytics for failure root cause detection in facilities management, Automation in Construction, 43(2014) 73-83, https://doi.org/10.1016/j.autcon.2014.03.012.

[64] Q. Lu, A. Parlikad, P. Woodall, G.D. Ranasinghe, X. Xie, Z. Liang, E. Konstantinou, J. Schooling, Developing a dynamic digital twin at building and city levels: A case study of the West Cambridge campus, ASCE Journal of Management in Engineering, https://doi.org/10.17863/CAM.45198.

[65] J.K.W. Wong, J. Ge, S.X. He, Digitisation in facilities management: A literature review and future research directions, Automation in Construction, 92 (2018) 312-326, https://doi.org/10.1016/j.autcon.2018.04.006.

[66] BuildingSMART, IFC 4 Officially Released, [Online] (12-03-2013). Available at: http://www.buildingsmart-tech.org/news/ifc4-officially-released2013.

[67] T. Liebich, IFC4—The New buildingSMART Standard, [Online]. Available at: http://www.buildingsmart-tech.org/specifications/ifc-releases/ifc4-release/buildingSMART_IFC4_Whatisnew.pdf.

[68] BSI 2014b, BS 1192-4:2014: Collaborative production of information Part 4: Fulfilling employer’s information exchange requirements using COBie – Code of practice, BSI Standards Limited.

[69] A. Tartakovsky, I. Nikiforov, M. Basseville, Sequential analysis: Hypothesis testing and changepoint detection, Chapman and Hall/CRC, 2014.

[70] N.R. Sakthivel, V. Sugumaran and S. Babudevasenapati, 2010. Vibration based fault diagnosis of monoblock centrifugal pump using decision tree. Expert Systems with Applications, 37(6), pp.4040-4049.

[71] R.P. Adams, D.J. MacKay, Bayesian online changepoint detection, 2007, arXiv preprint arXiv:0710.3742.

[72] D. Martínez-Rego, O. Fontenla-Romero, A. Alonso-Betanzos, J.C. Principe, Fault detection via recurrence time statistics and one-class classification, Pattern Recognition Letters, 84 (2016) 8-14, https://doi.org/10.1016/j.patrec.2016.07.019.

Table 1. Evaluation of IFC4 support for O&M management information requirements

	O&M Information Requirements	IFC4	COBie 2.4 (Spreadsheet xml)
Asset register information	Identification code/ unique reference/ barcode of asset	IfcIdentifier/ IfcGloballyUniqueId	Component sheet
	Description of asset	IfcLabel/IfcText	
	Status of asset	IfcLabel	
	Type of asset		
	Serial number	IfcIdentifier	
	Placement/location	IfcPlacement/ IfcSpace	
	Work manager, manufacturer, vendor	IfcPerson/IfcPerson AndOrganization	Type sheet
	Asset department	IfcOrganisation	
	Basic setting (e.g., output rating)	IfcLabel/IfcText	
	Category and code		
	Date of acquisition, installation or completion	IfcDateTime	Component sheet
	Permit-to-work requirement	IfcPermit	Job sheet
	Initial value, replacement cost, current value, disposal value, or written-down value	IfcCostValue	Job sheet
	Cost breakdown		
	Estimated Lifetime Remaining	IfcServiceLife	Type sheet
	Inspection or maintenance activity requirements	IfcTask/IfcEvent	Job sheet
	Inspection frequency and type	IfcTask	
	Other maintenance required	IfcTask/IfcEvent	
	Maintenance cost	IfcCostItem	
	Est. maintenance date, real maintenance date	IfcTaskTime	
	Contract code	IfcTask/IfcEvent	
	Accumulated depreciation		Spare sheet

	O&M Information Requirements	IFC4	COBie 2.4 (Spreadsheet xml)
Maintenance request related information	Placement/location	IfcPlacement/ IfcSpace	Job sheet
	Call number	IfcLabel	
	Call description	IfcText	
	Call details	IfcText	
	Assigned to which category		
	Person in charge	IfcPerson	
	Contact information	IfcPersonAnd Organization	
	Identification code of target asset	IfcIdentifier	
Sensor system information	Location identification	IfcPlacement/ IfcSpace	Component sheet Type sheet
	Location name	IfcLabel	
	Gateway identifier	IfcLabel	
	Gateway location	IfcPlacement/ IfcSpace	
	Timestamp of gateway	IfcTimeStamp	
	Gateway type	IfcLabel	
	Sensor identifier	IfcLabel	
	Gateway ID which sensor mapped to	IfcLabel	
	Sensor location	IfcPlacement/ IfcSpace	
	Identification code of target asset	IfcIdentifier	
	Asset name	IfcLabel	
	The type of sensor	IfcSensorType	
	Timestamp of sensor	IfcTimeStamp	
	Unit	IfcSensor	

	Source of components and spare parts				Description	IfcSensor	
	History record	IfcOwnerHistory/ IfcPerformanceHistory	Record in sheets		Value	IfcSensor	
	Risk related to people or property	IfcProperty EnumeratedValue		Space management information	Organisation identifier	IfcIdentifier	Floor sheet Space sheet Zone sheet
Building Management System	Site identifier	IfcSite	Facility sheet		Organisation name	IfcLabel	
	Site label	IfcSite			Site identifier and name	IfcSite	
	Node address/ Outstation number	IfcLabel	System sheet Component sheet		Building identifier and name	IfcBuilding	
	Outstation label	IfcLabel			Floor identifier and name	IfcBuilding Storey	
	Device response	IfcController			Room identifier and name, area	IfcSpace	
	Type of controller	IfcControllerTypeEnum			Room code	IfcSpace	
	Item label	IfcLabel			Occupancy activity, including identifier, occupier, occupancy time		
	Item units	IfcController					
	Power consumption	IfcTypeObjectProperty	Type sheet				
	Energy consumption and energy efficiency	IfcTypeObjectProperty			Record	IfcText	

* The empty block present that information is not defined in IFC schema; the grey texts present information is not defined completely for O&M management in IFC or COBie schema.