Received October 26, 2019, accepted November 25, 2019, date of publication December 6, 2019, date of current version December 19, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2958140

A Data-Driven Approach to Improve the Operation and Maintenance Management of Large Public Buildings

QIAO WEN[1], JIAN-PING ZHANG[1], ZHEN-ZHONG HU[1,2], XUE-SONG XIANG[3], AND TAO SHI[3]

[1]Department of Civil Engineering, Tsinghua University, Beijing 100084, China
[2]Tsinghua Shenzhen International Graduate School, Tsinghua University, Shenzhen 518055, China
[3]Kunming Chang Shui International Airport, Yunnan Airport Group Company, Ltd., Kunming 650211, China

Corresponding author: Zhen-Zhong Hu (huzhenzhong@tsinghua.edu.cn)

This work was supported in part by the National Key Research and Development Program of China under Grant 2018YFD1100900, in part by the National Natural Science Foundation of China under Grant 51778336, and in part by the Tsinghua University–Glodon Joint Research Center for Building Information Modeling (RCBIM).

ABSTRACT With the development of modern information technologies and more frequent utilization of information systems to operation and maintenance (O&M) management, a great amount of O&M data are collected nowadays. However, because of the large volume and poor quality, as well as a lack of effective data analysis techniques, these data are rarely analyzed and translated into useful knowledge for O&M decisions. This study presents a data model, which is named as datacube with multi-dimensional and unrestrained characteristics, for these data to better support data mining algorithms. The model organizes all the different data in both relational database and in the memories and is able to support analysis-requirements-oriented data extractions. Based on this datacube, an O&M data mining approach is proposed with procedures of data preparation, data clustering and data mining. The proposed datacube-based data mining approach was applied to the Kunming Chang Shui international airport terminal. More than 7 years on-site repairing data were used for data mining and the outcomes verified the model and the approach to be feasible and valuable for improving O&M management.

INDEX TERMS Operation and maintenance, data mining, datacube, airport terminal.

I. INTRODUCTION

Operation and maintenance (O&M) phase lasts the longest and costs the most within the building lifecycle [1]. With the development of modern information technologies and more frequent utilization of information system to facilitate facility management, a great amount of O&M data is collected nowadays. Therefore, researches have been focusing on the smoothly information handover from previous phases to the O&M stage [2], [3] and the analyzing and utilizing of O&M data [1], especially for large public buildings such as airport terminals, stadiums, convention centers, shopping malls, etc.

Some data-driven artificial intelligent methods were adopted to analysis the data collected during the O&M phase. For example, some studies applied data mining methods to analysis O&M data for an air-conditioning system in an educational building located in Montreal [4], to improve the energy efficiency for an international school campus in a tropical climate and an office building in a temperate European climate [5], and to optimize the geometrical, thermophysical and heating system attributes of apartments [6]. Such kind of methods are believed to have advantages in improving energy-saving behaviors and saving over 15% of the electricity used in the building [7]. Besides energy consumption analysis, analysis of the impact of occupants' behavior has been largely overlooked in building energy performance analysis [8]. Examples include a study that analyzed one-year observed data of an office building in Philadelphia to discover the occupancy schedule patterns and extrapolate the occupancy schedule [9], and a framework combining statistical analysis with 2 data-mining techniques, cluster analysis and association rules mining to identify valid window opening/closing operational patterns in the measured data of 16 offices in a natural ventilated office building in

The associate editor coordinating the review of this manuscript and approving it for publication was Qingchao Jiang.

Germany [10]. Moreover, office monitoring videos were captured and analyzed to detect occupancy [11] as well as occupant movement, light and equipment usages [12]. Lighting behavior were also considered useful for analyzing the overtime occupancy data [13] based on a stochastic model [14]. The raw data of all these studies are generated in Building Automation Systems (BAS), thus the data structures are relatively good for direct analysis. However, raw data with non-predefined structures are not included in this kind of analysis.

Besides performance optimization, automating the process of detecting equipment and system malfunctions to support proper diagnosis decision-making is also a hot area in both academic and industrial researches. The automated fault detection and diagnosis (FDD) methods consists of 3 mainstream methods [15]: process history-based, qualitative model-based and quantitative model-based methods. Process history-based method, which derives behavioral models from measurement data obtained from the process over time, is the only one to directly utilize building data among these 3 methods. Specifically, a two-stage back-propagation artificial neural network (BP-ANN) model combined with wavelet analysis and fuzzy logic was applied to FDD in an air handling unit (AHU) system [16]; a Bayesian network was used for AHU automatic FDD [17]; an ANN model based on the recursive deterministic perceptron was adopted to implement FDD at the whole-building level [18]. Other techniques have also been employed for FDD such as cluster analysis [19], fuzzy logic [20] and support vector machines (SVM) [21]. However, these methods only focus on a specific problem and thus the adopted data analysis methods are not generic for other operation and maintenance problems.

For some common facility management issues, data mining technology was utilized to analyze large data sets in the BAS of public buildings such as the tallest building in Hong Kong, with the aim of improving the building operational performance [22]. An optimization algorithm and computational fluid dynamics (CFD) simulations were coupled to improve the aerodynamic behavior of low-rise buildings with flat roofs [23]. Naïve Bayes, decision tree and SVM were also feasible to simplify the procedure of establishing models for predicting internal natural lighting and thermal comfort [24]. Computational intelligence and optimization approaches were used to improve a fuzzy controller's performance in an HVAC system, which aimed to find a method to moderate the energy use without compromising the comfort of the inhabitants [25]. BP-ANN algorithm was adopted to study the performance prediction of the ground source heat pump systems of an office building by real-time monitoring data and data-driven models [26]. Activity ontologies were developed and extended in order to capture flexible space-use patterns for user activities [27]. Nevertheless, these data mining methods need unique data pretreatment and analysis processes, and they usually depend on the embedded data models which are neither explicit to the public nor suitable to any kind of data sets.

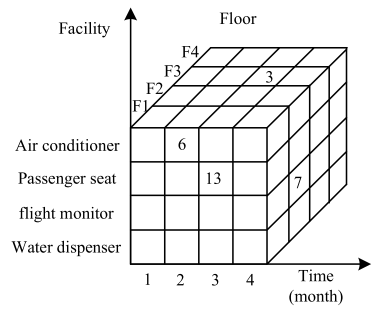

FIGURE 1. A 3-dimensional array example of a datacube for reported faults.

The above researches considered the building O&M data as the original data and applied various methods to achieve different research purposes. However, they also revealed some disadvantages considering current approaches to utilize data for decision support during O&M management of large public buildings, such as airport terminals which are complex in building composition, regular and frequent reparations, and always affecting a large number of passengers. For instance, the data were not well-organized that professional data managers should be involved to carry out data mining processes, the analysis processes were not closely related to the O&M management process, and there were no unique data analysis procedures to provide guidelines of dealing with that mess and large datasets. Consequently, it is necessary to develop more effective data analysis techniques to deal with these challenges [28].

This study presents a data model for O&M data to better support data mining algorithms. The model is named as datacube with multi-dimensional and unrestrained characteristics so that all the different data can be well-organized in both relational database and in the memories and be extracted according to analysis requirements dynamically. Based on this datacube, a hybrid O&M data mining approach is proposed with procedures of data preparation, data clustering and data mining. The proposed datacube-based data mining approach was applied to the Kunming Chang Shui international airport terminal. More than 7 years on-site repair data were used for data mining and the outcomes verified the model and the approach to be feasible and valuable for improving O&M management.

II. THE DATACUBE FOR O&M DATA OF LARGE PUBLIC BUILDINGS

A. THE DEFINITION OF THE DATACUBE

This study first presents a dynamic data-oriented information model named as datacube. A datacube, which is used to represent data along measure of interest, is an organized array of values, providing a more intuitive idea for data management and data mining. A datacube generally is a multi-dimensional concept which can be 3-dimensional or higher-dimensional. In any case, every dimension represents a separate measure whereas the cells in the cube represent the facts of interest.

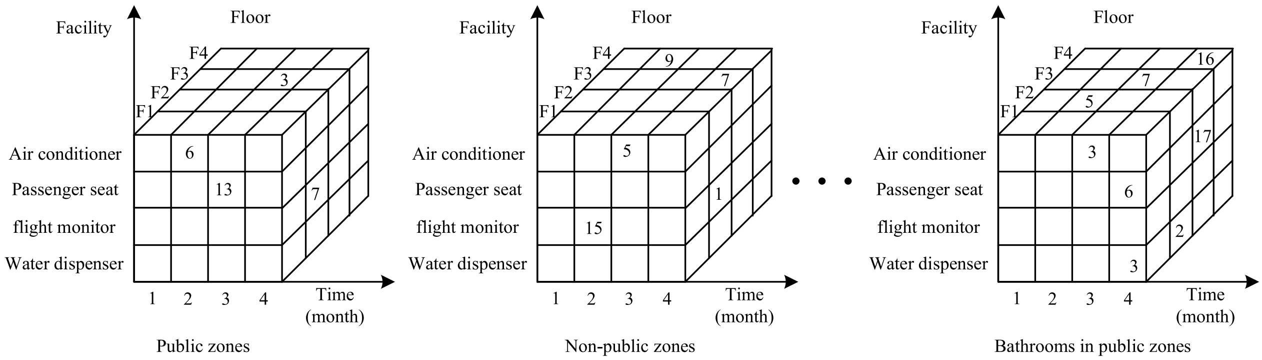

FIGURE 2. A 4-dimensional array example of datacubes for reported faults.

Sometimes cubes hold only few values with the rest being null.

Fig. 1 shows a datacube with 3 dimensions of time, facility, and location. The value of each cell represents the number of reported faults. Datacube models like this can also be represented by function shown in (1).

$$f : (X_1, X_2, \ldots, X_n) \rightarrow W \tag{1}$$

where X_i represents the i^{th} dimension, n represents the number of dimensions, and W is the value mapped to the cell. As shown in Fig. 1, n is 3, and X_1, X_2, X_3 are time dimension, facility dimension and floor dimension, respectively. The cell of number 13 indicates that there are 13 reported faults on the seats on the F1 floor in March, which can be expressed as f: (March, passenger seat, F1) → 13. Similarly, there are 7 monitor faults on the F2 floor in April, which can be expressed as f: (April, flight monitor, F2) → 7. The values of the unlabeled cells are null, representing no reported faults.

The dimension plays a key role in the datacube. A dimension is a structure that categorizes facts and measures. Typically, dimensions in a datacube are organized internally into one or more hierarchies. In practical applications, data analysis often requires more than 3 dimensions. Fig. 2 shows a 4-dimensional array after service dimension is added to the 3-dimensional array shown in Fig. 1. The dimensions can continue to increase to accommodate different decision-making requirements. For instance, "MEP sub-system", "fault type", etc. can be regarded as dimensions.

Decisions often require the collection of O&M data of multiple classes and levels of details. Through the flexible selection of dimensions and dimension hierarchies, the corresponding datacube can adapt to different measure levels and flexibly meet different demands of O&M data for various decision-makings.

B. INFORMATION COMPOSITIONS IN THE DATACUBE

1) OVERVIEW OF THE INFORMATION IN A DATACUBE FOR LARGE PUBLIC BUILDINGS

Due to the long time-span of the O&M period within a large public building, a huge amount of information is generated. Such information can be classified by format types, i.e., structured information like the name of a facility, the floor number; non-structured information like pictures and videos; and semi-structured information like some technical documents. It can also be classified by property types, i.e., facility information like doors, windows, seats; geometric information like geometric representations, location descriptions and coordinates; business information like operation tasks, regulations, technical manuals; and element information like volume, mass, material, price, etc. This information, if integrated in an information platform, can be stored in databases and transformed into a machine-readable datacube structure, by linking the unique identification (ID) through a union object.

Considering the complexity and variety of the information, combined with the practical O&M management requirements from the Kunming Chang Shui international airport terminal, a hierarchical structure is proposed to efficiently organize the information composition. The hierarchical structure is a tree structure that separates the managerial topics in different levels. Top levels focus on macro-scale areas or information, compared to bottom levels. This tree structure also facilitates the organizations of required information for data mining with various objectives in different data level. There's no strict reference for the number of levels or classification principles for each level while Fig. 3. shows the typical structure that was proposed for the Kunming Chang Shui international airport terminal project. It had three levels and an extra attribute level at the bottom so that the 3rd level nodes were connected with attributes of the corresponding facility in a machine-readable way, forming the standard facility data library.

2) KEY INFORMATION IN THE DATACUBE

Though there is a variety of data in the datacube, some of them play an important role in most of the data analysis scenarios. This section discusses these kinds of information.

a: LOCATION INFORMATION

According to the above-mentioned hierarchical structure, the location of a facility within a public building that covers a large area can be represented in a multi-scale way, i.e., floors,

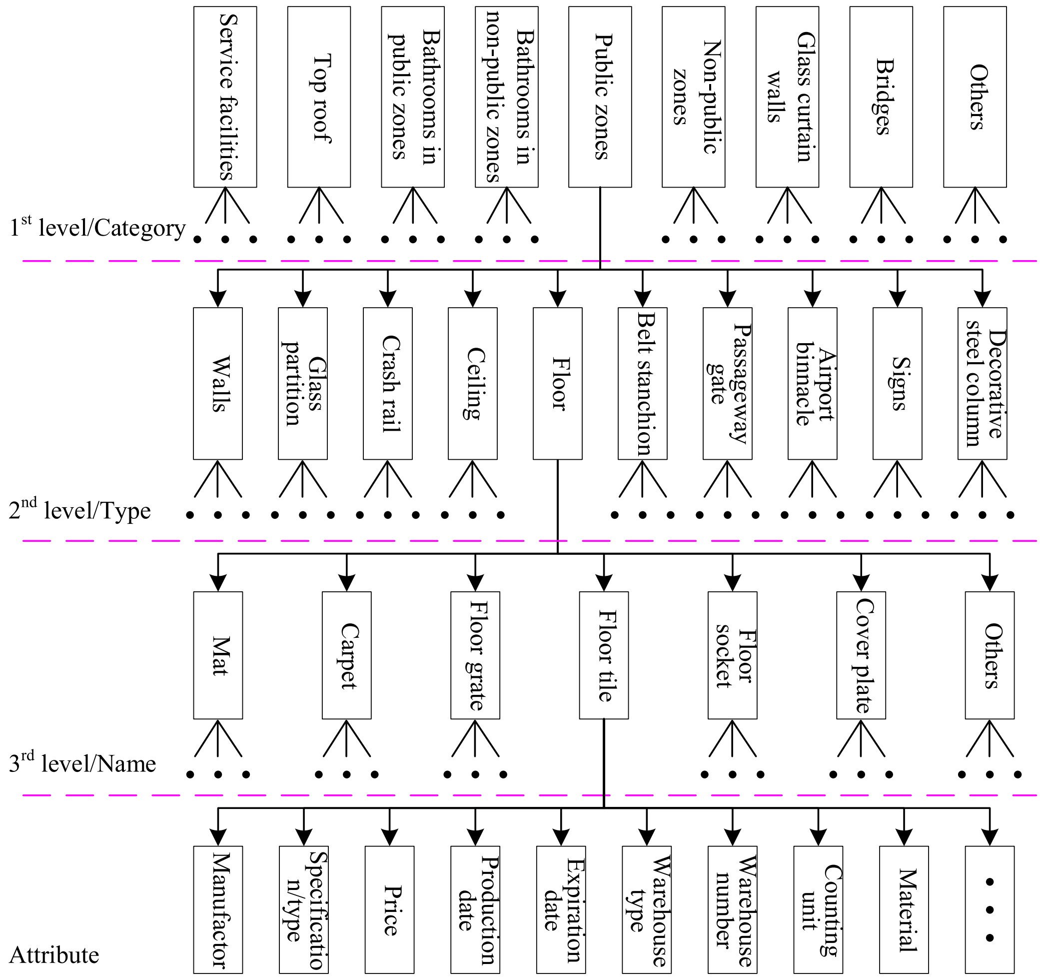

FIGURE 3. A typical structure for the O&M management of the Kunming Chang Shui international airport terminal project.

zones, areas and specific coordinates x, y and z, to facilitate daily management. The former three provide approximate positioning for a facility while the coordinates for precise positioning. Besides, the location classification can also follow the business process types, i.e., departure process zones and arrival process zones in an airport terminal.

b: PROPERTIES

The properties describe the status and characteristic of a large public building. They are the embedded information of a building itself. The properties are sometimes descriptive for managers which remains unchangeable for a long time or even permanently. For example, the material, price, factory, type and valid period of a facility; and the area, contractor, finish time, structure type and designed lifetime of the building structure. They can also be conclusive which may change through time by calculations according to the newly acquired information. For instance, the usage or distribution status of resources, the currency and cost caused by the operation tasks, and the workload and safe time period of the O&M management, etc.

c: SERVICE INFORMATION

The O&M management within a large public building usually involves multiple aspects of services such as property management, business management and cleaning/water supply, which can be used as the first level nodes of service data hierarchy according to the hierarchical structure. Then the specific tasks of each service can be regarded as the second level nodes. For example, a typical service data related to the repair task is shown in TABLE 1.

d: OTHER INFORMATION

Besides the above data, the daily O&M management also involves enterprise data, personnel data, etc. These data are also important to the daily O&M management. For example, if a cleaning contractor has not caused any serious accidents, the contractual relation will be maintained for a long period of time to avoid additional costs caused by re-bidding. Besides, the tasks are eventually finished by individual personnel whose data must also be included in standard human resources data library.

TABLE 1. Report repair data.

Name	Format	Explanation
Work order number	string	The unique number of the repair work order
Receive time	time	Manually inputted, accurate to the minute.
Contractor	string	Manually selected form standard enterprise data library with fixed format.
Facility	string	Manually selected form standard facility data library with fixed format.
Location/zone	string	Manually selected form standard location data library with fixed format.
Name	Format	Explanation
Fault type/cause	string	Partly manually selected form standard fault type/cause data library with fixed format; partly manual inputted in natural language.
Fault number	int	Manually inputted
Fault handling	string	Partly manually selected form standard fault handling data library with fixed format; partly manual inputted in natural language.
Fault photo	image	Manually uploaded
Arrive time	time	Manually inputted, accurate to the minute.
Finish photo	image	Manually uploaded
Finish time	time	Manually inputted, accurate to the minute.
Failed reason	string	Manual inputted in natural language.
Work order state	enum	In process state or shutdown state, and the default is in process.
Special repair work order number	string	If transferred to special repair, record the corresponding special repair work order number.
Spare parts entry/exit order number	string	If spare parts are involved, record the corresponding spare parts entry/exit order number.

3) THE DATA STRUCTURE FOR THE DATACUBE

In order to implement the datacube in either the database to storage, or the computer memory to facilitate the calculations, a relational data structure that supports arrays is proposed to store and manage the data via the star model. The star model contains a unique fact table, which stores the index IDs of multiple dimensions and the W values associated with them. Each index ID linked to a dimension table. Fig. 4 indicates the repair fact table that links to 6 dimension tables, i.e., facility, fault, location, time, contractor and service. Each dimension table contains different dimension hierarchies. For example, the time dimension table contains 4-dimensional hierarchies of day, month, quarter and year, while the floor dimension table contains only one hierarchy of the floor.

This structure can be a reference to design the relational database or the data structures to deal with the datacube. Specifically, in the relational database, via the IDs of the dimension tables, the repair fact table and each dimension table are joined so that the multi-dimensional data model shown in Fig. 1 or Fig. 2 is obtained. Herein, the number of repairs in each period can be obtained from the time dimension. Then by adding a facility dimension and a contractor dimension, the data for repairing different facilities and the amount of repairs of each facility by different contractors in each period can be obtained. It demonstrates that all different numbers of different repairs with multiple dimensions at different dimension hierarchies can be obtained through these combinations by adopting this data structure in a convenient way.

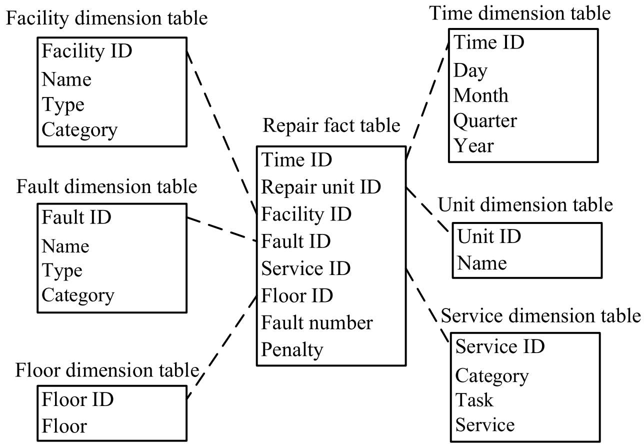

FIGURE 4. Star model of the repair topic.

III. A DATA MINING APPROACH BASED ON THE DATACUBE

This study proposes an approach for utilizing raw data generated during O&M management to improve the O&M management performance. The overview of the proposed approach includes 4 steps as illustrated in Fig. 5.

A. DATA PREPARATION

The first step is to integrate raw data from different sources, remove invalid data such as repeated data, and repair missing data or error data which are usually inevitable. Then according to the knowledge-based summarization, the data are carefully checked and revised before sending to the next steps.

A large amount of data are generated along with the O&M management progresses of large public buildings. During these processes, graphical data, standard libraries and operation processes are combined. Fig. 6 shows a typical data flow in a repair task. Specifically, when a facility problem is founded or sensed by the automation system, a repair process will be started, for example, triggered by a repair work order. Reported data in the work order usually include but not limited in the facility name, fault description, approximate location, and so on. Then a repair worker receives the work order and confirms the fault by selecting the facility from the standard data library or through the graphical interface. Once the facility is selected, the other related data can be

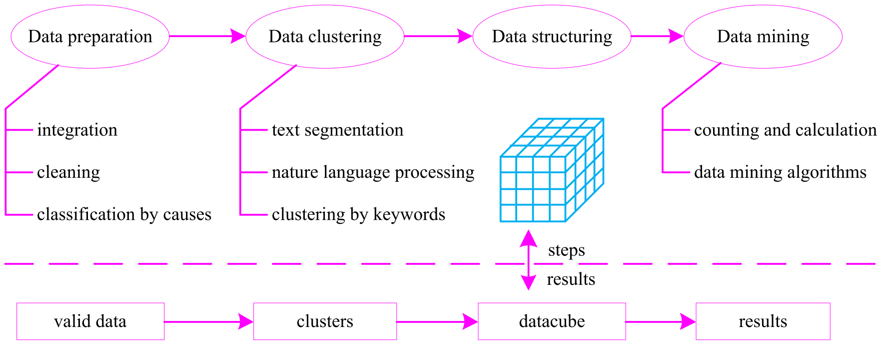

FIGURE 5. Methodology overview.

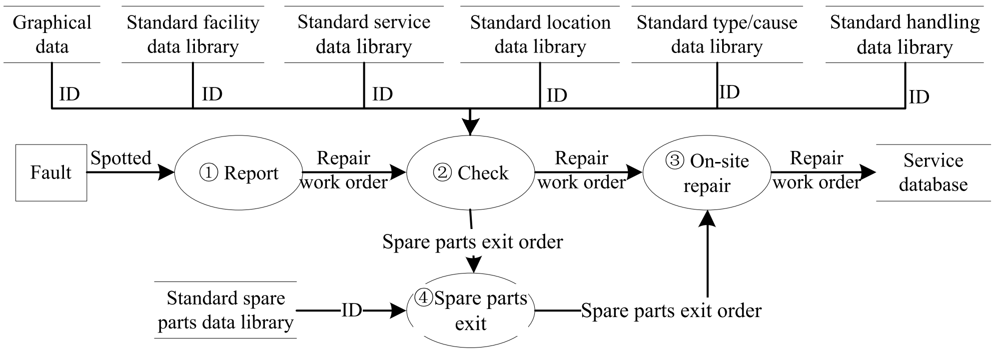

FIGURE 6. A typical data flow of the repair task.

retrieved from the database and appended to the repair work order. At the same time, if spare parts are required to fix the problem, an exit process for spare parts will be invoked at the same time. When the on-site work is finished, the results will be recorded in the database.

It can be found that the IDs of multiple standard libraries, instead of the specific task descriptions, are the most important data in the O&M management activities. Therefore, the datacube is suitable and capable to represent and integrate the information within these processes..

Particularly in this repair work example, the raw data scattered across multiple tables are first integrated into one data warehouse. Each repair report corresponds to multiple pieces of repair records with one-to-many relationships. Then data cleaning removes invalid data such as duplicated reports or invalid reports from the raw data to avoid interference. For instance, feedback records of repair work may include keywords such as ''false report'', ''repeated report'', ''no repair required'' or ''not within the scope of repair''. Such kinds of reports can be ignored in the succeeding data mining processes.

Data checking, which is the last step of preparing the data, helps to understand the records' meanings and ensures that each report corresponds to a certain facility. Necessary corrections can also be performed simultaneously. Take the repair work as the example, sometimes the reports are only recorded as ''fault repaired'' instead of recording the damage reason or the damaged part of the facility. Thus, this research first summarizes fault types by learning from repair records, i.e. the hand dryer faults include ''electronics faults'', ''hardware faults'' and ''unknown''. Furthermore, the classification can be multi-leveled, i.e., the ''electronics faults'' can be further divided to ''circuit element problem'', ''sensor problem'' and ''lock problem''. It facilitates workers to make an accurate record conveniently during both repair reporting and feedback after the task is completed. Finally, because of the various kinds of faults, it happens in some cases that there are many faults described similarly like ''water leakage'' or ''rain leakage''. It is necessary to analyze the causes of such ambiguous situations for accurate classification, as shown in TABLE 2.

B. DATA CLUSTERING

The data are then sent for clustering. It should be emphasized that the clusters can be used as the dimensions in the datacube. For example, when a series of clusters imply different locations or different facilities, they are considered as coordinates of location dimension or facility dimension.

Considering that the reports and feedbacks are usually recorded in natural languages, such as the location and fault descriptions in the report details and the feedback information, different with other clustering algorithms, this research first adopts text segmentation to obtain the keywords of the raw data, and then the combination of the keywords is

TABLE 2. Leakage classification.

Category	Subclass	Description
Water leakage	Structure leakage	Water leaks into civil structures, including walls, floors, canopies, etc.
	Condensate water	Condensate water due to the temperature difference, mainly HVAC system. It is a normal phenomenon.
	Water supply pipe leakage	Water supply network leaks, this leakage is a manifestation of damage to the water supply system
	Drainage pipe leakage	Rain and drainage network leaks, this leakage is a manifestation of damage to the rain and drainage system
	File pipe leakage	Fire pipe network leaks, the fire system is damaged when it leaks
Rain leakage	Top roof rain leakage	Top roof is exposed to the air, and when it leaks, it is damaged
	Bridge rain leakage	Boarding bridge is exposed to the air, and when it leaks, it is damaged

applied to cluster the raw date into different groups for further analyses.

Accordingly, the name and location of the damaged facility contained in the raw data can be considered as the keywords. The text segmentation, which is a part of natural language processing research [29], [30], usually involves 2 important parameters: dictionaries and stop words. Dictionaries consist of the meaningful words and are used to prevent segmenting meaningful words. For example, "boarding gate" is meaningful and "boarding" and "gate" both make sense too. This study stores words like "boarding gate" in the dictionary so that they will not be segmented apart. Furthermore, when a set of data has clear characteristics like a repair record must have one or more facility names, the locations, the results or comments, all these words are added to the dictionaries. Stop words usually consist of articles, prepositions, adverbs or conjunctions, etc., such as "a", "the" and "or". In this study, company names, staffs' names, digital numbers, single letters (A-Z) and some other specific words are considered as the stop words. Both dictionaries and stop words can be customized and improved via trials: run the segmentation process, check the results, update the dictionaries and stop words and then run the segmentation again with new dictionaries and stop words.

Location information is always important in a large public building. The location description can be semi structured data with a ranged description list of "boarding gate", "departure port", "arrival port", "parking space", "counter", "bathroom" and "room" in an airport terminal with some more detailed information in natural languages. Therefore, the data are clustered by region keywords with higher authority in this research. Then with requirements on more accurate analyses, further clustering can be conducted by the facility names or other characteristics.

C. DATA MINING

By now, the datacube can be generated through well-prepared data which are integrated, cleaned, checked and clustered. The dimensions for the datacube can be selected according to the clustering characteristics or any other selected properties related to facilities.

Then some intuitive results can be achieved via some simple counting and calculation, etc. But hidden patterns among the data need further data mining methods such as the Apriori or association rule mining algorithms. Specifically, an association rule refers to the form of $X => Y$ expression, where X denotes the premise and Y denotes the consequence. The 2 most important parameters of association rule mining are *support* and *confidence*. *Support* indicates the frequent occurrence of an association rule in all data, and *confidence* determines the frequency of the Y occurrence in all data with X included. Mathematically, *support* and *confidence* can be calculated by the probability, $P(X \cup Y)$, and conditional probability, $P(Y|X)$, as shown in (2) and (3), respectively,

$$Support(X => Y) = P(X \cup Y) \tag{2}$$

$$Confidence(X => Y) = P(Y|X) \tag{3}$$

where *support* is used for judging whether a rule is sporadic or not. Generally, a rule with *support* $\leq 10\%$ is considered sporadic, which means it has no practical meaning. *Confidence* is used for reliability of a rule: the higher the *confidence* is, the more reliable a rule is.

According to the clusters and the characteristics of large public building O&M records, this study presents a cluster-based pattern mining algorithm based on Apriori. In this algorithm, a hidden pattern is a logic implication that $F(A) => F(B)$ only when $F(A) \cup F(B) = \emptyset$. Here, $F(\cdot)$ represents a frequent status set and F(A), F(B) are the two sub sets of F(·). This logic implication means that if a record has all statuses in $F(A)$, it will have status in $F(B)$. Then, according to (2) and (3), a pair of $\{F(A), F(B)\}$ is considered as hidden pattern when it passes the following two tests of (4) and (5).

$$\begin{aligned} C(F(A) = > F(B)) &= P(F(B)|F(A))) \\ &= \frac{Num(F(A) \cup F(B))}{Num(F(A))} > C_{min} \end{aligned} \tag{4}$$

$$\begin{aligned} R(F(A), F(B)) &= \frac{P(F(A) \cup F(B))}{\sqrt{P(F(A)) \times P(F(B))}} \\ &= \frac{Num(F(A) \cup F(B))}{\sqrt{Num(F(A)) \times Num(F(B))}} > R_{min} \end{aligned} \tag{5}$$

where $P(\cdot)$ is the probability; C_{min} and R_{min} are given limitation by trials for acceptable results. For those pairs that passed the two tests, $F(A)$ is marked as the condition and $F(B)$ the consequence. The algorithm is described in pseudo-code as below.

Compared to classic Apriori algorithm, the proposed algorithm is approximately $N_p \cdot 2^{Np}/N_c$ times faster, where N_p is the total number of properties in the dataset and N_c is the number of clusters.

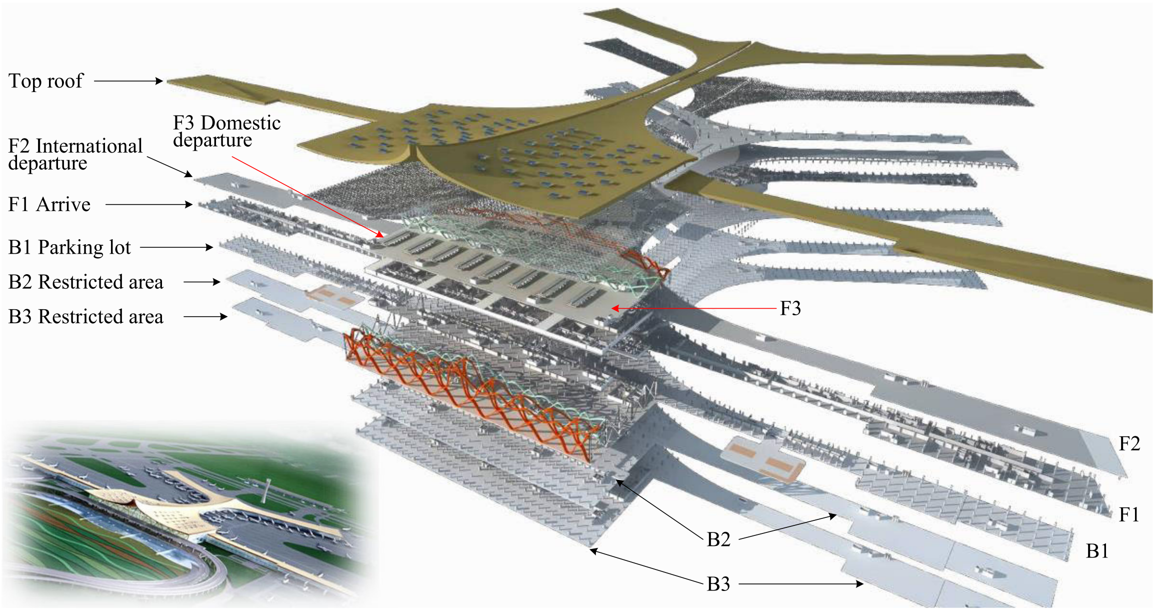

FIGURE 7. The 3D Model of the Kunming Chang Shui international airport terminal.

Algorithm 1 The Cluster-Based Pattern Mining Algorithm

Definition: N_{pi} is the number of properties in the i^{th} cluster; C_{min} is a pre-defined minimum count number; $(S_i)_j$ is the support count for the j^{th} property in the i^{th} cluster.

Foreach cluster S_i in S

1. Let $S_i = \emptyset$; $R = \emptyset$
2. Foreach $j = 1, 2, \ldots, N_{pi}$
   If $Count((S_i)_j) > C_{min}$ then
   $S_i = S_i \cup (S_i)_j$ // output S_i refers to the i^{th} cluster's longest frequent status sets
3. Foreach S_k in S_i
   If $Count(S_k)$ in all records $> C_{min}$ then
   $R = R \cup S_k$
4. Foreach $F(A)$ and $F(B)$ in F
   If pass tests (4) and (5) then
   $F(A) => F(B)$

IV. A CASE STUDY

A. OVERVIEW OF THE KUNMING CHANG SHUI INTERNATIONAL AIRPORT TERMINAL

The Kunming Chang Shui international airport is China's national gateway hub for Southeast Asia and South Asia and connects China to Europe and Asia. The total area of the terminal is 548,300 square meters, consisted of 4 floors above ground (F1, F2, F3 and F4) and 3 floors underground (B1, B2 and B3). Each floor is further divided into 8 different zones of A-H. At the same time, the terminal is divided into departure process zones and arrival process zones. Fig. 7 is the 3D model of the terminal, with an aerial view on the lower left corner. It officially opened on June 28, 2012. The passenger throughput in 2017 was 44.73 million [31].

Each floor has two wings on the front, except F3. B3 and B2 are restricted areas with lots of tunnels and secret facilities. B1 is a parking lot. F1 is the arrival for both domestic and international flights; while F2 and F3 are for international and domestic departure, respectively. Between F3 and top roof, there is a small catering area existed labeled F4. Due to the enormous passenger throughput, the facilities were occasionally damaged. The O&M management department built a daily report and repair system in the same year as the airport opened and the system has been working till now for the unified reporting, disposal and cancellation of the daily report and repair of all facility faults. The raw data were accumulated from then on.

B. WORKFLOW OF THE REPAIR TASK

Fig. 8 shows the workflow of the repair task, it consists of 2 parts: report process as shown in Fig. 8 (a) and repair process as shown in Fig. 8 (b).

The report process includes only 1 step, that is, employee A reports the fault information of the facility, including location description (*location_description*), fault details (*report_details*), etc., and the process ends.

The repair process consists of the following 7 steps:

Step ①: Employee B logs into the system and clicks on the new repair report. In this step, the system automatically records the employee's name and the receiving time (*receive_time*, accurate to the second).

Step ②: Employee B manually determines whether the repair report is valid or not. Invalid repair reports include repeated repair reports, false reports, repair reports that are

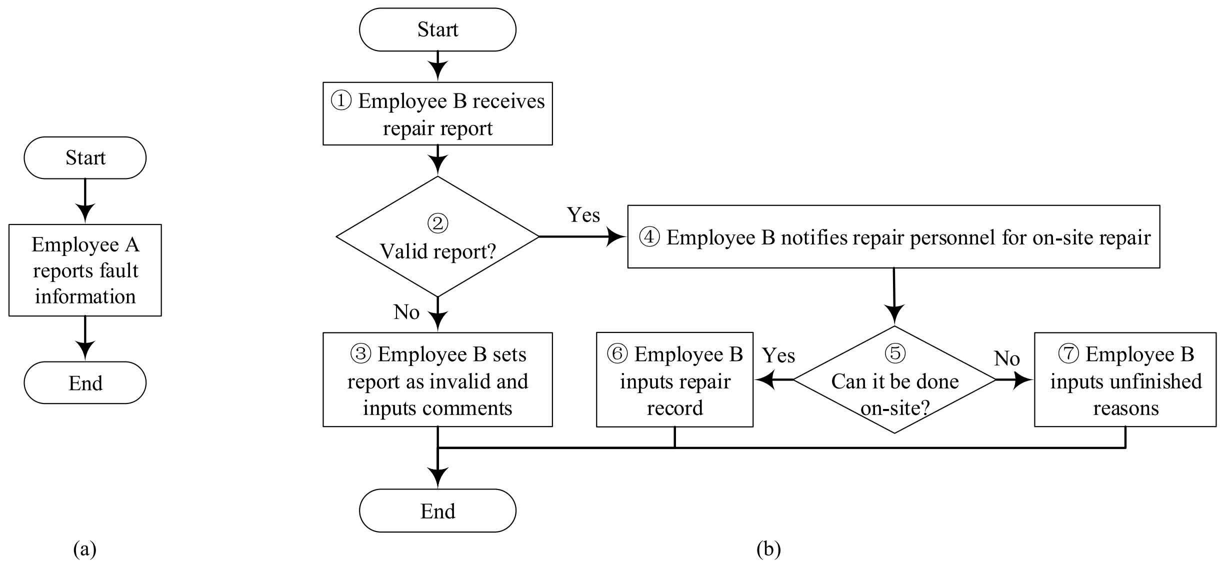

FIGURE 8. Report process and repair process.

not within the scope of airport responsibility and repair reports that have to be transferred to special repair task.

Step ③: If the repair report is invalid, employee B inputs the corresponding feedback information (*feedback_record*) to the system, close the repair task to end the process. Such information has no fixed format, and the employee B records it in natural language and saves it to the feedback information (*feedback_record*).

Step ④: If the report is valid, employee B immediately notifies the repair worker to the site.

Step ⑤: The repair worker goes the site and judges whether it can be mended on-site or not. The arrival time (*arrive_time*, accurate to the minute) must be recorded no matter what the judgement result is.

Step ⑥: If fault can be repaired on-site, the repair work starts immediately. The repair record will be reported. There may be multiple pieces of repair records corresponding to possible multiple repairs. After the repair work is finished, employee B will complete the final repair record with time information (*finish_time*, accurate to the minute). The repair record data and the repair final result data are also not in a fixed format but recorded by employee B in natural language and saved to the repair record (*repair_records*).

Step ⑦: If it cannot be repaired on-site, employee B should record the unsolved reason as the final result. Non-repairable situations include reports of a false alarm, or repairs that are too difficult to be solved on-site and should be transferred to monthly or quarterly professional repair schedule, namely the special repair task. The recording process is the same as step ③ and the feedback (*feedback_record*) are saved in the form of natural language.

C. DATA ACQUISITION

The information system mentioned above was designed to be as simple as possible, so that it could be used by anyone

TABLE 3. Region keywords and corresponding report numbers.

Restroom	Room	Boarding gate	Counter	Misc.
7883	6037	4026	1855	1722
Parking space	Arrival port	Departure port	Top roof	F4
1151	737	581	45	31

TABLE 4. Top 10 region-facility clusters and corresponding report numbers.

Cluster	Number	Cluster	Number
Door – restroom	1006	Paper towel dispenser - restroom	718
Door - room	985	Pit latrine - restroom	713
Water dispenser - boarding gate	962	Toilet - restroom	615
Door lock - room	751	Door - boarding gate	612
Door lock - restroom	731	Stall door - restroom	581

without too much training. It was a B/S (Browser/Server) website and showed only a few pages for users to input, search and edit their reports. It was officially launched on July 5, 2012 thus had been running smoothly for over 7 years, accumulating 24,068 repair reports. Fig. 9 illustrates some of the raw data of the F3 floor.

The raw data were greatly different in format and content. A total number of 148 different facilities were obtained in the form of a table with 22,146 rows and 148 columns, after a total number of 1,922 invalid repair reports were cleaned up during data preparation. TABLE 5 is a summary of the terminal's facility faults and serves as the initial reference for developing the periodical on-site inspection program.

During the clustering process, 10 regional clusters (see TABLE 3) and further 567 region-facility clusters were obtained (the top 10 clusters by frequency are shown in TABLE 4).

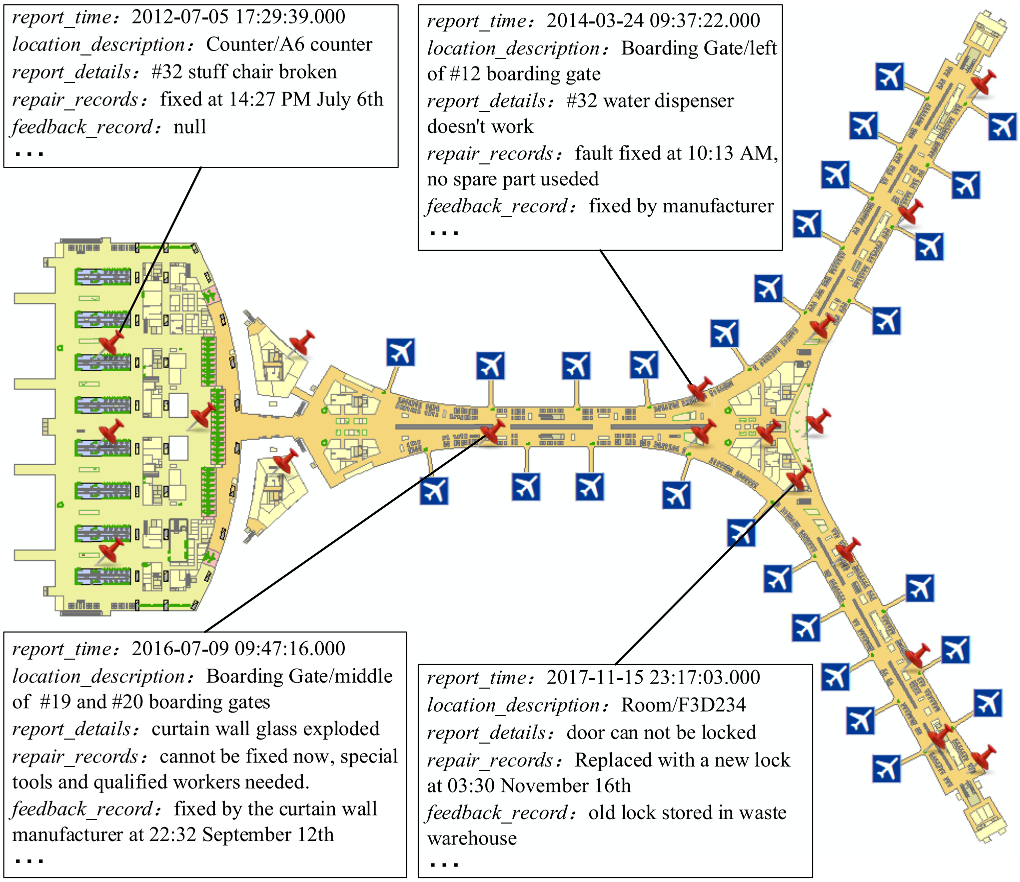

FIGURE 9. Raw data examples.

D. GENERATING THE DATACUBE

After the processes of data preparation and data clustering, the above-mentioned raw data were organized in datacube-based relational databases. According to different data mining objectives, different datacube were dynamical generated to proceed data analysis algorithms. Fig. 10 shows a general datacube generated according to different facilities with three dimensions: time, floor, and facility.

Then every cell in the cube can be represented as a function "f:(Facility_ID, Month_ID, Floor_ID)-> Fault numbers" or as a multi-dimensional array with vectors in the form of "(Facility_ID, Month_ID, Floor_ID, Fault numbers)". According to Fig. 10, when "floor = F1" and "month = April", there were 4 records including (Air conditioner, April, F1, 7), (Passenger seat, April, F1, 1), (Flight monitor, April, F1, 4) and (Water dispenser, April, F1, 5).

E. RESULTS AND DISCUSSIONS

After generating the datacube, data analysis and data mining were conducted to find useful and meaningful patterns for management. The following sections show some managerial inspirations raised by such kind of data-driven management approach.

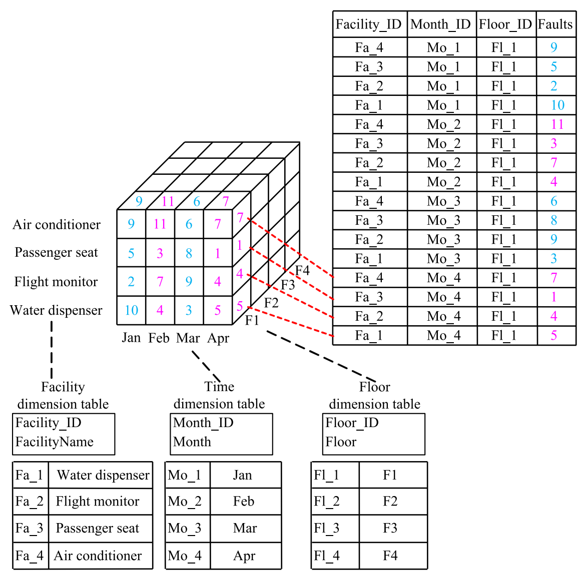

FIGURE 10. A general datacube generated according to different facilities.

1) FACILITY PRIORITY RANKS

A feasible periodical on-site inspection is expected to have limited number of priorities: it is unrealistic to set a priority for every facility. Thus, it is necessary to rank the facilities

TABLE 5. Facility fault hierarchy of Kunming Chang Shui international airport terminal.

1st level	2nd level	3rd level / facility (frequency)				
terminal	component	structure water leakage (264)	top roof (393)	wall (102)	ceiling (441)	other (112)
building	floor	floor grate (64)	floor transition strip (18)	mat (24)	carpet (11)	cover plate (31)
		antistatic floor (455)	floor tile (86)	rubber floor (59)		
	door	passageway gate (313)	passageway gate knob (125)		passageway gate lock (74)	
		room door (551)	room doorknob (224)		room door lock (464)	
	glass	curtain wall glass (152)	partition glass (133)	door glass (137)	canopy glass (12)	glass breaker (14)
water	water supply	water supply pipe leakage (504)		no water (48)		
supply	sewer	drainage pipe (62)	grease interceptor (10)			
and	fire control	fire emergency lamp (3)	smoke detector (4)	fire alarm (8)	fire shutter (22)	safety exit (126)
drainage		fire hydrant cabinet (20)	file pipe leakage (16)	fire sprinkler (2)	fire control switch (2)	fire hose (2)
HVAC	HVAC	supply and exhaust duct (80)	condensate water (30)	air conditioner (99)		
electric	power supply	electricity (1008)	socket (156)	floor socket (150)	power box (13)	
	lighting	lamp (1030)	lamp shell (15)			
	weak power	digital newspaper device (6)	LED screen (138)	television (83)	public telephone (89)	intercom (62)
	elevator	moving walkway (129)	elevator (194)			
public	sign	sign (81)				
service	railing	crash rail (542)	belt stanchion (80)			
	water dispenser	water dispenser (1504)	paper cup holder (81)			
	smoking room	door (20)	doorknob (6)	lock (9)	cigarette lighter (122)	
	other	mobile charging station (10)	airport seat (74)	vending machine (5)	children's play facility (14)	
bathroom	male, female and handicap bathroom	hat-and-coat hook (32)	pit latrine (1088)	hand dryer (245)	urinal (90)	mirror (24)
		bathroom entrance doorknob (35)	paper towel dispenser (813)	shower accessories (11)	toilet (592)	soap dispenser (466)
		handicap bathroom door (439)	stall door lock (728)	stall door (1092)	trash can (48)	bathroom tile (302)
	baby care room	door (28)	doorknob (8)	door lock (62)	sofa (21)	
	cleaning room	beverage water heater (116)	mop bucket (123)	door (23)	doorknob (10)	door lock (68)
	common	piping shaft gate lock (27)	sink (446)	faucet (416)	floor drain (75)	
civil	security check	security-check hardware (8)	security-check gate (84)	cabinet (50)	luggage inspection system (10)	
aviation		security-check gate lock (240)	computer camera (8)	security-check gate knob (109)		
	check-in	luggage inspection room door (18)	luggage inspection room doorknob (9)	luggage inspection room door lock (102)	printer (235)	counter hardware (96)
		luggage check-in system (205)	self-check-in system (21)	luggage scale (39)	X-ray security machine (97)	
	boarding	boarding gate lock (32)	boarding gate knob (58)	boarding gate (504)	broadcast (102)	airport binnacle (22)
	arrival	arrival gate (310)	arrival gate knob (67)	arrival gate lock (18)	luggage reclaim system (29)	
	departure	terminal gate (160)	terminal gate knob (27)	terminal gate lock (10)		
	boarding bridge	gate (500)	gate knob (139)	gate lock (13)	other instrument (170)	
		side gate (224)	side gate knob (128)	side gate lock (12)		
	VIP	room door (131)	room doorknob (50)	room door lock (16)		
	common	surveillance software (13)	surveillance camera (13)	computer (89)	network system (45)	fax machine (3)
		flight monitor (307)	flight query system (17)	clock (49)	flight departure control system (8)	
		scanning gun (234)	desk-chair (222)	display (55)	identity authentication system (101)	
Misc.	Misc.	pest control (5)	room cabinet (11)	projector (2)		

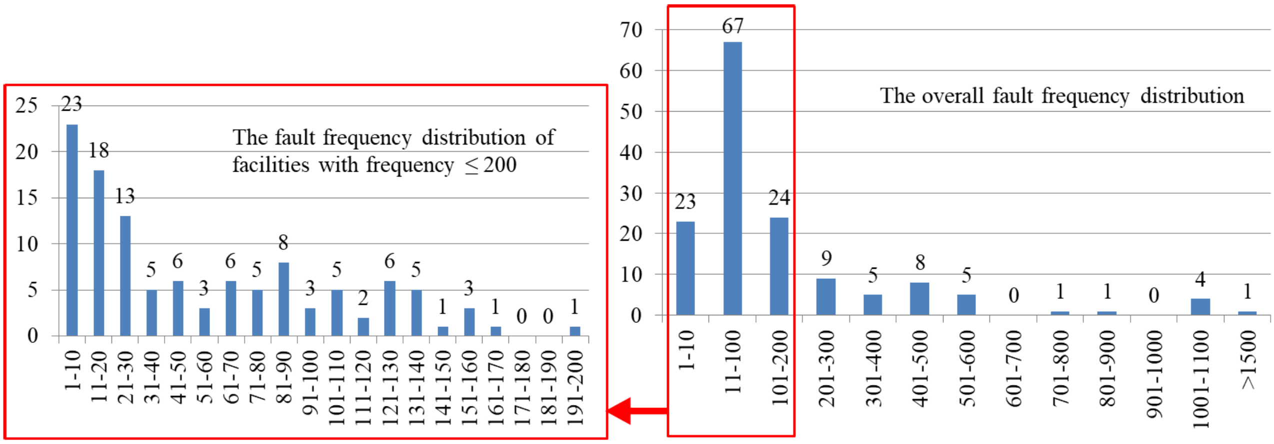

FIGURE 11. Facility fault frequency distribution.

TABLE 6. The relationships between facilities and their frequencies.

Frequency interval	≤ 30	31 - 170	191 - 600	≥ 601	Total
Facility number	54	59	28	7	148
Proportion	36.5%	39.9%	18.9%	4.7%	100%
Summary of frequencies	741	5,337	10,232	7,263	23,573
Proportion	3.1%	22.6%	43.4%	30.8%	100%

According to the fault frequencies shown in TABLE 5, a facility fault frequency distribution chart is obtained, as shown in Fig. 11. The vertical axis represents the number of facilities, and the horizontal axis stands for the frequency interval. It can be found that the number of facilities with frequencies less than 200 is 114, accounting for 77% of the total number (148), which means that a majority facility is usually in good condition for a long period of time, and the facility intact rate is high for the terminal. By magnifying the reports of facilities with frequencies no larger than 200 times, it can be found that the number of facilities with frequencies that were no larger than 30 was 54, accounting for 36.5% of the total number (148); the number of facilities with frequencies between 30-170 was 59, accounting for 39.9% of the total.

Therefore, the discrete frequencies were divided into 4 intervals as shown in TABLE 6. The table indicates that the number of facilities with frequencies that are no greater than 30 is 54, accounting for 36.5% of the total number (148), but the sum of their frequencies is only 3.1% of the total sum (23,573). The number of facilities with frequencies greater than 600 is 7, accounting for 4.7% of the total, but the sum of their frequencies is 30.8% of the total. It indicates that the main source of the repair reports is only a small part of the often-damaged facilities, while most only fail occasionally. This conclusion proves that a single priority for all facilities is infeasible, the priorities should be decided according to the fault frequency ranks.

At the same time, a discussion also argued that the facilities, which can cause serious consequences once they failed, must be added to the first rank. Such facilities include luggage inspection system, luggage check-in system, X-ray security machine, luggage reclaim system, network system, flight query system, flight departure control system and identity authentication system, which were manually assigned. The other 3 ranks remain unchanged. As a result, the priority ranks were suggested as shown in TABLE 7, which involves 4 ranks referring to the 4 fault frequency intervals. This rank is the foundation for deciding the following important things for the periodical on-site inspection program.

TABLE 7. Facility priority ranks.

Ranks	1st class	2nd class	3rd class	4th class
Standard	frequency≥ 601 or high importance	191 ≤ frequency ≤ 600	31 ≤ frequency ≤ 170	frequency ≤ 30
Facility number	15	27	56	50

2) INSPECTION PERIODS & SPARE PARTS

The best periods for each rank group should be the average time since last failure. The facility at the middle of each frequency interval can be considered as the representative of each rank group. However, the O&M managers proposed another way. They needed a fixed schedule for employees to follow easily with weekend excluded. On the other hand, both the O&M management department and outsourcing companies should prepare adequate spare parts for good periodical on-site inspection in an airport terminal. As a result, according to the data mining results, The O&M managers were suggested some rules for scheduling inspection periods as well as storing spare parts in the terminal as shown in TABLE 8. Confirmed by the O&M managers, such rules did help saving inspection costs and better utilizing the warehouse spaces inside the terminal.

3) STRATEGY TO IMPROVE FACILITY DISTRIBUTIONS

According to the O&M management department, after the official opening on June 28, 2012, the airport staff and facilities had a running-in period that resulted in many repair reports at the beginning. The O&M management became stable after several months. Thus, some data between January,

TABLE 8. Suggested rules for inspection and spare parts.

Ranks	Period	Instructions for inspection	Instructions for storing spare parts on site	Fund priorities
1st class	A day	Carry out at 9 a.m. every working day	Always prepare at least 5 pieces for each facility	Priority
2nd class	A week	Carry out every Monday after the daily inspection	Always prepare at least 2 pieces for each facility	Normal
3rd class	A month	Carry out at the 1st working day of each month after daily inspection	Can be in shortage for no more than one week for each facility	Normal
4th class	A quarter	A full coverage inspection at the first working day of each quarter	Can be in shortage for no more than two weeks for each facility	Normal

2013 and December, 2018 were used for the pattern mining. Some typical results are shown in TABLE 9 where the minimum *support* is set to 0.1 with the minimum *confidence* equals to 0.5.

As shown in TABLE 6, 35 facilities' frequencies are no less than 190. The minimum *support* is set to 0.1, which means that nearly only this part of the data is merely included. Another important parameter is a minimum *confidence* = 0.5. Then, pattern mining is done.

12 patterns are founded, with the highest probability of which are bathroom-related facility => water dispenser and the patterns between bathroom-related facilities. The second pattern is of no meaning because they are located together and with the same usage. The first pattern, with the consequence of water dispenser, is as shown in TABLE 9. Pattern 1 has the highest *support* of all found patterns. Other patterns with *support* less than 0.1 are not included.

The O&M managers confirmed that all the above 3 patterns are reasonable. Water dispensers require water sources, so their installation locations are all outside bathrooms with no exception. That is, all water dispensers are adjacent to a bathroom, but not every bathroom has a water dispenser outside. These 3 patterns also indicate that passengers usually like to drink some water after using bathrooms. According to this inference, the study suggested setting up water dispensers outside as many bathrooms as possible to meet passengers' habits.

Besides the above-mentioned findings, a related research also utilized data mining algorithms to find some hidden rules that were helpful for FDD of the mechanical, electrical and plumbing (MEP) systems [1] in the same project. Application results show that both these two researches benefit the management department, however, this study presents the idea of datacube as a more common data model for data mining algorithms, thus is more generic to different projects and various optimization objectives.

V. FINAL REMARKS

Building O&M data generated and collected during management have potentials to improve O&M performance. However, it is usually hard to take advantages of such data due to their varieties in formats, contents and meanings, etc., especially when they are not translated into useful knowledge. This study presents a datacube as a data model with multi-dimensional and unrestrained characteristics. The model organizes all the different data in both relational database and in the memories and is able to support analysis-requirements-oriented data extractions. Based on this datacube, a hybrid data mining approach is proposed with procedures of data preparation, data clustering and data mining.

TABLE 9. Patterns between bathroom-related facilities and water dispenser.

No.	Rules	Support	Confidence
Pattern 1	Pit latrine + urinal + toilet => water dispenser	0.237	0.570
Pattern 2	Stall door => water dispenser	0.227	0.606
Pattern 3	Pit latrine + urinal + toilet, stall door => water dispenser	0.111	0.642

The proposed datacube idea and the correspondent approach were applied to the facility repair task of Kunming Chang Shui international airport terminal from July 2012. More than 7 years on-site repair data counts up to over 24,000 records were used for data mining. Some of the raw data were structured stored in database while some were recorded in natural language form. With the help of the proposed approach, the priority ranks were suggested for different facility groups, some rules for inspection and spare parts were established, as well as some hidden relationships such as bathroom and water dispenser were found by association rule mining.

However, the idea of datacube is not yet well developed in theoretic way with more evidence for popularity, for example, how to extract more meaningful data in an efficient way with convincible mathematical prototype is remained to be further studied. Other efforts such as building up a knowledge base according to the mining results and introducing ontology or knowledge graph to obtain a more accurate way to carry out data mining algorithms will also be the focus points in future.

REFERENCES

[1] Y. Peng, J.-R. Lin, J.-P. Zhang, and Z.-Z. Hu, "A hybrid data mining approach on BIM-based building operation and maintenance," *Building Environ.*, vol. 126, pp. 483–495, Dec. 2017, doi: 10.1016/j.buildenv.2017.09.030.

[2] S. Rodriguez-Trejo, "Hierarchy based information requirements for sustainable operations of buildings in Qatar," *Sustain. Cities Soc.*, vol. 32, pp. 435–448, Jul. 2017, doi: 10.1016/j.scs.2017.03.005.

[3] J. Patacas *et al.*, "BIM for facilities management: Evaluating BIM standards in asset register creation and service life planning," *J. Inf. Technol. Construct.*, vol. 20, no. 10, pp. 313–318, 2015.

[4] Z. Yu, F. Haghighat, B. C. M. Fung, and L. Zhou, "A novel methodology for knowledge discovery through mining associations between building operational data," *Energy Buildings*, vol. 47, pp. 430–440, Apr. 2012, doi: 10.1016/j.enbuild.2011.12.018.

[5] C. Miller, Z. Nagy, and A. Schlueter, "Automated daily pattern filtering of measured building performance data," *Autom. Construct.*, vol. 49, pp. 1–17, Jan. 2015, doi: 10.1016/j.autcon.2014.09.004.

[6] C. Alfonso, S. Gianluca, P. M. Savino, and G. Daniele, "Data mining for energy analysis of a large data set of flats," *Proc. Inst. Civil Eng.-Eng. Sustainability*, vol. 170, no. 1, pp. 3–18, 2017, doi: 10.1680/jensu.15.00051.

[7] E. Azar and C. C. Menassa, "Agent-based modeling of occupants and their impact on energy use in commercial buildings," *J. Comput. Civil Eng.*, vol. 26, no. 4, pp. 506–518, 2012, doi: 10.1061/(ASCE)CP.1943-5487.0000158.

[8] E. Delzendeh, S. Wu, A. Lee, and Y. Zhou, "The impact of occupants's behaviours on building energy analysis: A research review," *Renew. Sustain. Energy Rev.*, vol. 80, pp. 1061–1071, Dec. 2017, doi: 10.1016/j.rser.2017.05.264.

[9] X. Liang, T. Hong, and G. Q. Shen, "Occupancy data analytics and prediction: A case study," *Building Environ.*, vol. 102, pp. 179–192, Jun. 2016, doi: 10.1016/j.buildenv.2016.03.027.

[10] S. D'Oca and T. Hong, "A data-mining approach to discover patterns of window opening and closing behavior in offices," *Building Environ.*, vol. 82, pp. 726–739, Dec. 2014, doi: 10.1016/j.buildenv.2014.10.021.

[11] J. Zou, Q. Zhao, W. Yang, and F. Wang, "Occupancy detection in the office by analyzing surveillance videos and its application to building energy conservation," *Energy Buildings*, vol. 152, pp. 385–398, Oct. 2017, doi: 10.1016/j.enbuild.2017.07.064.

[12] T. Zhang, P.-O. Siebers, and U. Aickelin, "Modelling electricity consumption in office buildings: An agent based approach," *Energy Buildings*, vol. 43, no. 10, pp. 2882–2892, Oct. 2011, doi: 10.1016/j.enbuild.2011.07.007.

[13] K. Sun, D. Yan, T. Hong, and S. Guo, "Stochastic modeling of overtime occupancy and its application in building energy simulation and calibration," *Building Environ.*, vol. 79, pp. 1–12, Sep. 2014, doi: 10.1016/j.buildenv.2014.04.030.

[14] X. Zhou, D. Yan, T. Hong, and X. Ren, "Data analysis and stochastic modeling of lighting energy use in large office buildings in China," *Energy Buildings*, vol. 86, pp. 275–287, Jan. 2015, doi: 10.1016/j.enbuild.2014.09.071.

[15] W. Kim and S. Katipamula, "A review of fault detection and diagnostics methods for building systems," *Sci. Technol. Built Environ.*, vol. 24, no. 1, pp. 3–21, Jan. 2018, doi: 10.1080/23744731.2017.1318008.

[16] B. Fan, Z. Du, X. Jin, X. Yang, and Y. Guo, "A hybrid FDD strategy for local system of AHU based on artificial neural network and wavelet analysis," *Building Environ.*, vol. 45, no. 12, pp. 2698–2708, Dec. 2010, doi: 10.1016/j.buildenv.2010.05.031.

[17] M. Najafi, D. M. Auslander, P. Haves, and M. D. Sohn, "A statistical pattern analysis framework for rooftop unit diagnostics," *HVAC&R Res.*, vol. 18, no. 3, pp. 406–416, Jun. 2012, doi: 10.1080/10789669.2011.607881.

[18] F. Magoulés, H.-X. Zhao, and D. Elizondo, "Development of an RDP neural network for building energy consumption fault detection and diagnosis," *Energy Buildings*, vol. 62, pp. 133–138, Jul. 2013, doi: 10.1016/j.enbuild.2013.02.050.

[19] I. Khan, A. Capozzoli, S. P. Corgnati, and T. Cerquitelli, "Fault detection analysis of building energy consumption using data mining techniques," *Energy Procedia*, vol. 42, pp. 557–566, Jan. 2013, doi: 10.1016/j.egypro.2013.11.057.

[20] F. Lauro, "Building fan coil electric consumption analysis with fuzzy approaches for fault detection and diagnosis," *Energy Procedia*, vol. 62, pp. 411–420, Jan. 2014, doi: 10.1016/j.egypro.2014.12.403.

[21] T. Mulumba, A. Afshari, K. Yan, W. Shen, and L. K. Norford, "Robust model-based fault diagnosis for air handling units," *Energy Buildings*, vol. 86, pp. 698–707, Jan. 2015, doi: 10.1016/j.enbuild.2014.10.069.

[22] F. Xiao and C. Fan, "Data mining in building automation system for improving building operational performance," *Energy Buildings*, vol. 75, pp. 109–118, Jun. 2014, doi: 10.1016/j.enbuild.2014.02.005.

[23] Y. Qiu, B. San, and Y. Zhao, "Numerical simulation and optimization of wind effects of porous parapets on low-rise buildings with flat roofs," *Adv. Civil Eng.*, vol. 2019, p. 11, 2019, Art no. 3402613, doi: 10.1155/2019/3402613.

[24] A. Ahmed, N. E. Korres, J. Ploennigs, H. Elhadi, and K. Menzel, "Mining building performance data for energy-efficient operation," *Adv. Eng. Informat.*, vol. 25, no. 2, pp. 341–354, Apr. 2011, doi: 10.1016/j.aei.2010.10.002.

[25] S. Hussain, H. A. Gabbar, D. Bondarenko, F. Musharavati, and S. Pokharel, "Comfort-based fuzzy control optimization for energy conservation in HVAC systems," *Control Eng. Pract.*, vol. 32, pp. 172–182, Nov. 2014, doi: 10.1016/j.conengprac.2014.08.007.

[26] L. Yan, "The performance prediction of ground source heat pump system based on monitoring data and data mining technology," *Energy Buildings*, vol. 127, pp. 1085–1095, Sep. 2016, doi: 10.1016/j.enbuild.2016.06.055.

[27] X. Chen, T. W. Kim, J. Chen, B. Xue, and W. Jeong, "Ontology-based representations of user activity and flexible space information: Towards an automated space-use analysis in buildings," *Adv. Civil Eng.*, vol. 2019, p. 15, 2019, Art no. 3690419, doi: 10.1155/2019/3690419.

[28] Z. Yu, B. C. M. Fung, and F. Haghighat, "Extracting knowledge from building-related data—A data mining framework," *Building Simul.*, vol. 6, no. 2, pp. 207–222, Jun. 2013, doi: 10.1007/s12273-013-0117-8.

[29] NLP group of Computing Institute of Chinese Academy of Sciences. (2018). *NLPIR Chinese Lexical Analysis System*. Accessed: Nov. 11, 2018. [Online]. Available: http://ictclas.nlpir.org/

[30] The Stanford Natural Language Processing Group. (2018). *Stanford Word Segmenter*. Accessed: Nov. 26, 2018. [Online]. Available: https://nlp.stanford.edu/software/segmenter.html

[31] Civil Aviation Administration of China. (2018). *Statistical Data of All Airports in China*. Accessed: Jun. 11, 2018. [Online]. Available: http://www.caac.gov.cn/

QIAO WEN received the B.S. degree in municipal engineering and the M.S. degree in civil engineering from the Harbin Institute of Technology, China, in 2009 and 2012, respectively, and the Ph.D. degree in civil engineering from Tsinghua University, China, in 2019.

His research interests are building information modeling (BIM), information technologies in construction industry, and intelligent information processing technology.

JIAN-PING ZHANG received the B.S. degree from Tsinghua University, China, in 1980, and the Ph.D. degree from The Hong Kong Polytechnic University, in 1996.

She is currently a Professor with the Department of Civil Engineering, Tsinghua University. Her research interests focus on CAD/CAE technology, 4D-CAD, BIM, intelligent information processing technology, and digital disaster prevention and intelligent decision.

ZHEN-ZHONG HU was born in Guangdong, China. He received the B.S. and Ph.D. degrees from the Department of Civil Engineering, Tsinghua University, China, in 2005 and 2009, respectively.

He is currently an Associate Professor with the Tsinghua Shenzhen International Graduate School and the Department of Civil Engineering, Tsinghua University. His research interests focus on information technology in civil engineering, BIM, and digital disaster prevention and mitigation.

XUE-SONG XIANG was born in Yichang, Hubei, China, in 1972. He received the B.S. degree in electronic engineering (signal and information processing) from the Xi'an University of Electronic Science and Technology, China, in 1994, and the M.S. degree in aviation safety management from the National Civil Aviation School, France, and the National Higher School of Aeronautical Manufacturing Engineers, France, in 2003.

He is currently the Vice Manager of Kunming Chang Shui International Airport. He is interested in project management and facility management.

TAO SHI was born in Hunan, China, in 1985. He received the B.S. degree in heating ventilation and air conditioning (HVAC) from the Huazhong University of Science and Technology, Wuhan, China, in 2008.

He joined Kunming Chang Shui International Airport, in July 2010. Up to now, he has been engaged in the allocation of room resources, data statistics, and operation and maintenance management of the terminal. He is also an experienced expert in operation and maintenance management of airport terminals.