RESOLVNG INCONSISTENCY IN BUILDING INFORMATION USING UNCERTAIN KNOWLEDGE GRAPHS: A CASE OF BUILDING SPACE MANAGEMENT

Xiang Xie[1], Janet Chang[2], Mohamad Kassem[1] and Ajith Parlikad[2]
[1]School of Engineering, Newcastle University, Newcastle upon Tyne, UK
[2]Institute for Manufacturing, University of Cambridge, Cambridge, UK

Abstract

The challenge of inconsistent and conflicting information in building space management is addressed using the uncertain knowledge graph. In this method, facts are represented in first-order logic formulas with weights attached to quantify credibility. Conflict-free information is acquired through the Maximum A-Posterior query over the Markov Logic Network. Semi-structured interviews with asset management professionals helped understand the general space management process in commercial buildings and its challenge of conflicting information. A case study showed that trusted space information can be drawn from dissimilar sources by eliminating redundant and inconsistent information, hence, reducing expensive manual verification processes.

Introduction

The construction industry is known for its fragmented and project-based nature. Yet, different processes are generating a large volume of data in disparate formats as a building evolves through the design, construction, operation, maintenance and decommissioning phases of the building lifecycle. For example, architects and engineers generate design data during the planning and design phase of a project. Then, contractors of varied disciplines produce a wide variety of asset data to meet the building owner's asset information requirement while incorporating any changes that occur to the project during the construction phase. Moreover, a set of as-built information about a building, also known as handover information, is transferred to an asset owner upon project completion for supporting various post-construction activities. Handover information is used for varieties of post-construction activities and is continuously updated as the building receives several types of capital and recurring maintenance. Given that the average life of a commercial building is about 50 years, one building could generate a significant volume of data and information during its life (Spire, 2022).

Buildings have been progressively complex, and this trend in the building industry leads to generating a significant volume of handover information in diverse formats. Accordingly, multiple asset information management systems become inevitable for managing varied types of handover information. For example, the building owner can use Computer-Aided Facility Management (CAFM), Computerized Maintenance Management System (CMMS), Integrated Workplace Management System (IWMS), Energy Management System (EMS), and Electronic Document Management System (EDMS) for managing data with distinctive features. Building Information Modeling (BIM) intends to promote the interoperability of fragmented information (ISO/IEC, 2015), which may be embedded in design drawings, Construction Operations Building Information Exchange (COBie) spreadsheets, Operation and Maintenance (O&M) manuals, and many other documents. To complement and enhance the BIM method, semantic web technologies have shown their strength in improving data exchange processes, linking data stemming from diverse domains, and allowing complex data querying and reasoning (Pauwels et al., 2017). Focused on BIM Level 3, the Information Management Framework (IMF) sets out the guidelines for technical standards, processes, and interoperability frameworks to integrate and utilize multi-source machine-interpretable data (Hetherington and West, 2020).

When managing a portfolio of assets, inconsistent processes of information flow from varied systems often lead to ambiguities or contradictions in data that hinder informing decisions based on practical knowledge. In the building industry, data duplication and redundancy are regularly found in diverse domains, phases, and software because a wide range of stakeholders generates data in various stages without managing the accumulated data (Fang et al., 2022; Zadeh et al., 2017). For example, in addition to handover information, recurring building asset surveys collect incremental (e.g., deterioration) or radical (e.g., alteration or complete renovation) changes that may be stored in systems such as CAFM. Although BIM can be used as a single information repository, the handover information and operation data are commonly managed separately due to technological barriers such as poor compatibility between BIM and other systems (Durdyev et al., 2022). The individually managed systems may contain associated but contradictory information, which makes it impossible to determine the right and accurate information to use. This frequently triggers resource-intensive processes of verifying the collected information. This paper presents an uncertain knowledge graph-based method to determine conflict-free building information by eliminating the potential inconsistencies in different representations of information.

This paper utilizes the space management process of commercial buildings to verify the feasibility of the proposed method. Space management involves maximizing the use of space while promoting the productivity and well-being of people using space. Therefore, effective space management depends on reliable space availability, its intended use, utilization status, etc. (Becerik-Gerber et al., 2012). Additionally, the information used to verify the space assignment often comes from different departments in the organization. For example, the estate division provides the necessary plans and spatial information. The designated departments provide the space assignment, which can be validated through finance records for space use. Moreover, the Human Resources (HR) employee directory is repeatedly used for employee verification. To accommodate the information, many organizations use labor-intensive processes for managing the space regardless of the advancements in BIM. This manual exercise is expensive and time-consuming, especially for those who manage multiple buildings. Therefore, the space information from diverse sources (i.e., estates, HR) is encoded to First-Order Logic (FOL) in a Markov Logic Network (MLN) with weights indicating the credibility, and the Maximum A-Posterior (MAP) reasoning is conducted to obtain the most likely space utilization condition based on the knowledge graph.

The remainder of the paper is organized as follows. The literature review presents the conventional space management process and the associated information requirements. In addition, the usage of knowledge graphs in building information management is discussed. In the methodology section, the space management process is summarized, in which the uncertain knowledge graph is used to resolve conflicts using MAP inference. A case study is presented to validate the proposed method followed by the discussion and conclusion.

Literature review

Space management

Space management is about optimizing the use of space to accommodate the delivery of organizational objectives (Azizi et al., 2020). The primary goal of space management is to improve space use efficiency, meaning maximize the use of space while minimizing the relevant operating costs. According to Rogers (2002), space planning and relevant data analysis are prerequisites for increasing efficiency in managing space. Accurate information about space, combining space inventory, utilization, and continual space management, is critical to support these requirements (Abdullah et al., 2013). Space inventory is to identify the availability of space. Measuring space utilization based on the frequency and occupancy of the space is required to understand the effective use of the space. Managing space is not limited to evaluating space requirements but also includes future use of the space. Azizi et al. (2020) argue that strategic space planning needs are achievable with accurate information about space use. Further, updated information about space prevents space shortage and unnecessary capital investment for expansion and operation expenditure on under-utilized space.

Despite the vital role of space management in organizations, the literature evaluating essential information needed to optimize space use is scarce. Liu et al. (1994) assert that handover information such as floor plans and design specifications are critical information for operation support, but the study fails to describe the use of floor plans for space management clearly. In recent years, several studies propose using occupancy-sensing technologies to maximize the use of space. For example, Azizi et al. (2020) propose utilizing occupancy and booking information to optimize space and energy use in higher education institutes. Valks et al. (2018) present the use of real-time space-tracking tools to search for an available space to accommodate the needs of the end-users. These studies prove the effectiveness of occupancy sensing to gain the efficiency of commonly used meeting rooms, but such tools are impractical to track the space used for other types of space, such as individual offices because of privacy issues. To contribute to this gap and support the technical development proposed in this paper, semi-structured interviews are conducted with asset management professionals to understand the existing space management processes to assess the utilization of all types of space and acquire the relevant information that is essential for verifying and validating the space assignment.

Knowledge graph for building information management

Knowledge graphs (KGs), also known as semantic networks, are widely adopted as an efficient approach to representing information to derive new knowledge. In the past few years, KGs have shown considerable progress in automated fraud detection, investment advice chatbots, and market intelligence. By describing the objective world's concepts, entities, and relationships in the format of graphs, KGs can structure, manage, and understand the information that emulates human cognitive thinking. In the Architecture, Engineering, and Construction (AEC) industry, heterogeneous data is dispersed across siloed databases or systems. Thus, KG can demonstrate its strength in establishing a graphical knowledge base, which allows to integrate structured or unstructured information, enabling complex queries and reasoning for the extraction and inference of required knowledge (Pauwels et al., 2021).

The exploitation of KGs comprises three stages: (1) KG construction, (2) KG deduction, and (3) KG usage (Li et al., 2021). In the KG construction stage, a set of interconnected types of entities, their relationships and attributes are created for information representation, with a single or federated ontology as the schema defining the domain vocabulary to be used. The ontology is a formally written representation of a certain domain of interest, agreed upon and shared among a group of people. A list of commonly used ontologies in the AEC industry is summarized in Pauwels et al. (2021), including those related to building topology modeling (e.g., Building Topology Ontology, BOT for short) and those concerning building Mechanical, Electrical, and Plumbing (MEP) systems (e.g., Brick and Haystack ontologies). While most efforts are put into the KG construction through the definition of appropriate domain ontologies and delivering existing knowledge, relatively fewer studies have been conducted on the synthesis and deduction of new knowledge based on known facts. Examples of KG deduction are rule checking and regulation compliance checking. Using logic-based inference engines, extra information can be automatically inferred from the existing information in Resource Description Framework (RDF) and Web Ontology Language (OWL) formats based on a set of custom rules. The arranging KG interacts with users in an orderly manner ensuring prompt access to needed knowledge.

Despite the wide adoption of KGs in the AEC industry, their applications involving temporal building information are scarce. This is mainly due to the uncertainty or inconsistency in the building knowledge representations for the following reasons: (1) the contradiction of ontology, which may include ill-defined concepts, relations, etc.; (2) the repetition involved in collecting similar building information; (3) the presence of diverse information sources; and (4) the use of inaccurate information (Anand and Kumar, 2022). Apart from the first challenge associated with developing an ontology, all the other challenges that affect uncertainty are very pertinent to the AEC sector. Therefore, inherent uncertainties of facts need to be included in developing the KG with confidence scores (Ji et al., 2021). Aligning this assumption, the most probable building information can be estimated using the MAP inference, which integrates complementary information and eliminates potential inconsistencies.

This paper aims to showcase the use of uncertain KGs in integrating multi-source information for space management purposes. One of the important reasons why labor-intensive and manual processes still dominate in space management is that the information is scattered across different sources with potential conflict with each other. Information retrieval, error correction and verification heavily rely on professionals' experiences. With the capabilities of representing facts residing in different sources and deducting most likely facts with higher credibility, uncertain KGs are tested for the very first time.

Methodology

To assess the feasibility of the proposed method, this paper summarizes the 'real-world' space management processes. The reason for selecting space management is because the management process is simple and uses predominately graphical handover information such as site, floor and furniture plans, together with various non-handover information. Contrary to existing studies, this paper aims to leverage space-related handover and non-handover information from multiple sources and resolve potential information duplications and redundancy for the space management processes in the commercial building sector.

Space management process

Collecting a deeper understanding and insights into the existing space management processes requires a rich and real-world context that cannot be emulated in an artificial setting. Therefore, semi-structured interviews are conducted with nine asset management professionals who are involved in managing spaces within their organizations. Each organization manages a wide range of commercial buildings with varied use. Participants' quotes and phrases describing the space management process with the supporting information were captured. Then, the collected data were translated into a diagram for evaluation (see Figure 1). After comparing nine individual processes, the common space management process was extracted, as well as the essential information needed, the steps of using the requisite information, and decision points in the process. This is used as evidence for the case study. The interviews were recorded and transcribed, with the permission of the participants to enhance analysis.

The interviews revealed that space management exercises are semi- or annual events for most of the participants in order to meet the organizational business objectives. A mixture of the site plan, floor layout, and room schedules is widely used, in addition to furniture plans, and non-spatial information such as the employee directory, space assignments, and financial records for confirming the space use. Due to the lack of automated information update processes, the information about space availability usually is not up-to-date, and the space use verification process is labor-intensive and time-consuming. For example, the participant who manages over 400 buildings reported that it takes up to six months to validate the existing space. In managing space, a survey is sent to the existing occupants for their verification. Then, the survey results are compared with the previously submitted survey results, the existing space assignments, employee verification or payroll records. Site verification is conducted annually, and additional validation is necessary when there are significant differences between the previous survey and the newly submitted one or disputes among the occupants. Therefore, uncertain knowledge graphs are introduced as a tool to accommodate information in different representations and eliminate conflicts among multiple information from disparate information systems that may arise while managing the space. This effort would further reduce resource-intensive verification tasks.

Maximum a-posteriori inference for uncertain KG

In this paper, KGs are assumed to be encoded in the RDF format. The RDF model can be seen as a graph where each RDF triple (s, p, o) is represented as a node-edge-node structure. In other words, a tripe is represented as an edge between the subject s and the object o labeled by the predicate p. The Markov Logic Network (MLN) is used to extend the RDF graph with uncertainties (Richardson and Domingos, 2006). MLN combines the Markov network with the First-Order Logic (FOL) by attaching a weight to each first-order formula and the weighted formulas can be seen as the template of the Markov network.

First-order logic has been proven to be a general, flexible, and powerful framework for knowledge representation and reasoning. In practice, RDF triples can be automatically converted to logical statements in first-order logic through

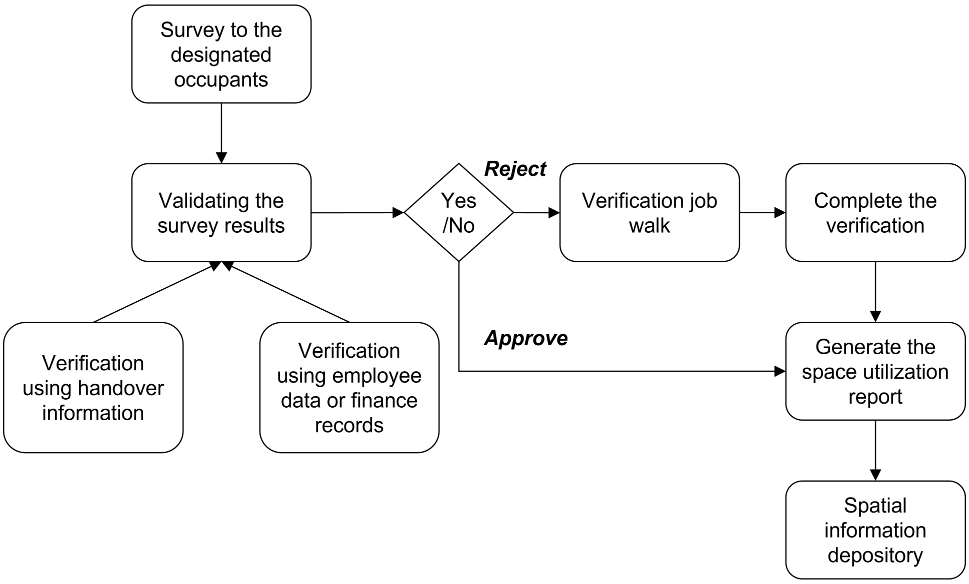

Figure 1: Diagram of information flow for space management

SPARQL query (He et al., 2014). The RDF knowledge graph is equivalent to a finite set of implicitly conjoined ground atoms. A ground atom denotes an atomic formula that applies a predicate to a tuple of constants. For example, O ranges over offices in one building, and D is the department which occupies the building. The ground atom $hasSite(O,D)$ indicates that office O is owned by or in other scenarios allocated to department D, which is either true or false. Formulas are constructed by recursively connecting atomic formulas using logical connectives (e.g., negation, conjunction, disjunction, implication, and equivalence) and quantifiers (e.g., universal quantifier and existential quantifier). For automated inference, it is common to convert formulas to the clausal form (conjunctive normal form, CNF for short), which is a disjunction of ground atoms. However, the FOL is limited to hard facts and cannot express uncertainty. The KG becomes completely impossible as long as a single inconsistency is found. Therefore, the MLN is introduced by attaching weights to the FOL formulas, in which way the knowledge is represented in a probabilistic sense.

An MLN M_{LC} is a finite set of pairs (F_i, w_i), where F_i is a logic formula in the clausal form and w_i is the weight suggesting the probability or credibility of the fact modeled. A log-linear probability distribution is defined to describe the possibility of the world $\mathbf{x}$:

$$P(\mathbf{X}=\mathbf{x}) = \frac{1}{Z}\exp(\sum_{i=1}^{F} w_i n_i(\mathbf{x})) \tag{1}$$

in which F is the number of formulas in M_{LC}, and $n_i(\mathbf{x})$ is the number of true groundings of F_i in $\mathbf{x}$. The most probable and conflict-free world is obtained by solving the maximization problem given evidence $\mathbf{E}=\mathbf{e}$:

$$\arg\max P(\mathbf{X}=\mathbf{x}|\mathbf{E}=\mathbf{e}) \tag{2}$$

This is also known as Maximum A-Posterior (MAP) inference. The maximization problem can be transformed into an integer linear programming (ILP) problem. The ILP variables $x_l = \{0,1\}$ are associated with each ground atom occurring in each clause F_i.

Let $L^+(F_i)$ denote the set of ground atoms occurring unnegated in F_i, while $L^-(F_i)$ denotes those occurring negated in F_i. For every single F_i, the linear constraints introduced into the ILP include:

$$\sum_{l\in L^+(F_i)} x_l + \sum_{l\in L^-(F_i)} (1-x_l) \geq z_i, \text{ if } w_i \geq 0 \tag{3}$$

$$\begin{aligned}\sum_{l\in L^+(F_i)} x_l + \sum_{l\in L^-(F_i)} (1-x_l) \\ \leq (|L^+(F_i)| + |L^-(F_i)|)z_i, \\ \text{if } w_i < 0\end{aligned} \tag{4}$$

where $z_i \in \{0,1\}$ is the binary variable associated with ground clause F_i. For hard facts with weight $w_i = \infty$, the linear constraint becomes:

$$\sum_{l\in L^+(F_i)} x_l + \sum_{l\in L^-(F_i)} (1-x_l) \geq 1, \text{ if } w_i = \infty \tag{5}$$

Finally, the objective function of the ILP is:

$$\max \sum_i w_i z_i \tag{6}$$

The ILP is very established and there exist efficient ILP solvers (Richardson and Domingos, 2006). In summary, the information required for implementing space management comes from multiple systems, and the knowledge contained in these dissimilar information systems can be viewed as a single large formula or conjunction of clauses/formulas. The knowledge stored is usually timestamped. According to the time when the knowledge is generated and the characteristics of acquired knowledge, weights can be affiliated with each clause implying the likelihood of such facts. The most likely and conflict-free information is queried by solving the corresponding ILP problem. The MAP information can be understood as the combination of the most reasonable parts of individual knowledge sources, after carefully corroborating with each other. The MAP inference based on uncertain KGs is illustrated in Figure 2.

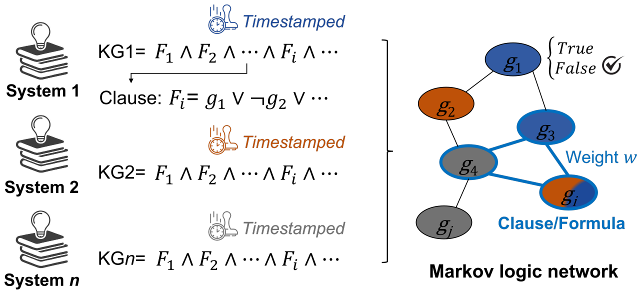

Figure 2: MAP inference based on multi-source information

Case study

In 2021, the Association of University Directors of Estates (AUDE) annual publication reported a substantial budget increase is intricately connected to the size of the university estates. The expenditure associated with operating space is the second highest cost next to the staff salary. Therefore, it is in the best interest of each university to maximize the use of the existing spaces while reducing the relevant costs. In the UK, Higher Education Institutes (HEIs) are mandated to report the use of their estates, including the space utilization of teaching, research, and non-educational spaces, as part of the space measurement to the Higher Education Statistics Agency (HESA) for benchmarking and justifying future investment for expansions. In response to this mandated requirement, each university has established its own processes for managing space suitable for the size of estates with a slight variation of information used. However, the interview results suggested that the core space management process of managing space is similar. Therefore, this case study utilizes the common process extracted from the interviews to demonstrate the process, together with the information flow from various sources.

An example used in this paper demonstrates the process of determining the most likely space utilization by integrating information from diverse sources with potential conflicts. The case study assumes that assorted handover information such as site plans, building layouts, room schedules, and furniture plans are used to confirm the spatial information. This information originates from the handover information depository. For verification purposes, the employee directory or payroll records are used to confirm the space assignment.

A building *bldg* owned by University *Uni* is shared by two departments *Dept_A* and *Dept_B*. Three spaces are used to describe this process, *Office_A* to *Office_C*, providing workplaces for all staff employed by the two departments (*Staff_A* to *Staff_D* here). Figure 3 shows the RDF graph of the space utilization-related information, coming from the handover information and the human resources management system. Two ontologies are used in this case, Building Topology Ontology (`https://w3c-lbd-cg.github.io/bot/`) and The Organization Ontology (`https://www.w3.org/TR/vocab-org/`). As shown in the figure, *Office_A* and *Office_B* were assigned to *Dept_A* according to the design, and the others were assigned to *Dept_B*. Information regarding the staff working in the University is recorded in the HR system, including the workspace assigned to each of them and their affiliations.

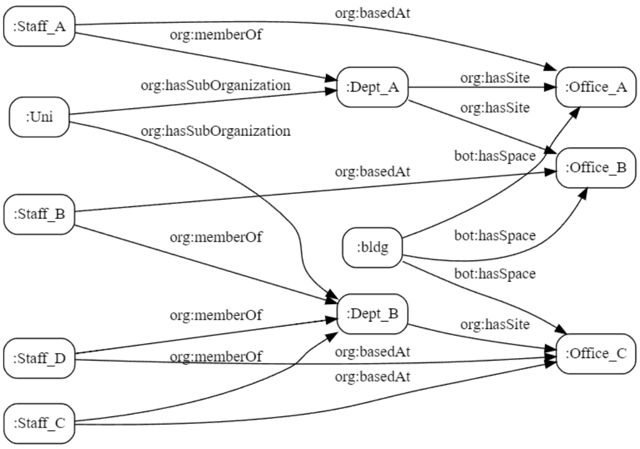

Figure 3: RDF graph of the space utilization information

An obvious conflict can be found. *Office_B* was assigned to *Dept_A* (*Dept_A hasSite Office_B*) according to the design records. However, *Staff_B* who is affiliated with *Dept_B* (*Staff_B memberOf Dept_B*), occupies the workplace *Office_B* (*Staff_B basedAt Office_B*). To resolve conflicts like this, the inherent uncertainties of facts lying in different pieces of information need to be quantified. For simplicity, we assume the weights of facts are inversely proportional to the duration between the current time and its creation. The older the more likely it is affected by inaccuracy information. Besides, three constraints are considered:

$$\begin{array}{ll}
(1) & memberOf(x,y) \wedge basedAt(x,z) \Leftrightarrow hasSite(y,z) \\
(2) & \forall x \exists y : hasSite(x,c), \neg hasSite(y,c) \\
(3) & \forall x \exists y : memberOf(c,x), \neg memberOf(c,y)
\end{array}$$

The second and third constraints can be hard constraints with infinite weights if there are clear regulations that one working space cannot be shared by two departments and each staff cannot have a dual affiliation. These constraints are converted to CNF for the benefit of inference:

$$\begin{array}{lcl}
(1) & (\neg memberOf(x,y) \vee \neg baseAt(x,z) \vee & w_1 \\
 & hasSite(y,z)) \wedge & \\
 & (\neg hasSite(y,z) \vee memberOf(x,y)) \wedge & \\
 & (\neg hasSite(y,z) \vee baseAt(x,z)) & \\
(2) & \neg hasSite(x,c) \vee \neg hasSite(y,c) & w2 \\
(3) & \neg memberOf(c,x) \vee \neg memberof(c,y) & w3
\end{array}$$

Corresponding weights are given to each fact (evidence) and constraint. In this case, the evidence from the design records is given a weight of 0.8 ($Dept\_A\ hasSite\ Office\_B$), and the evidence from the HR system is given a weight of 1.6 ($Staff\_B\ memberOf\ Dept\_B$ and $Staff\_B\ basedAt\ Office\_B$). Constraint (2) and (3) are assumed to be hard constraints ($w_2, w_3 = \infty$). For constraint (1), because Markov logic only assigns weights only to disjunctions of clauses, the weight associated with the first constraint is equally divided into three clauses and the clauses are split into individual formulas. Figure 4 shows the structure of MLN, in which a node indicates the grounding fact that it is a Boolean random variable, and a clique represents a logic formula (for example, the three red lines correspond to the first constraint).

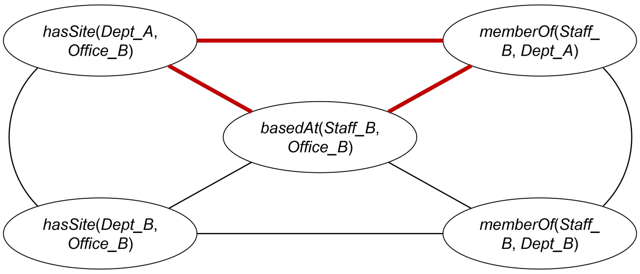

Figure 4: MLN of the space utilization and its groundings

Based on the MLN and associated weights, the MAP inference is conducted by solving the corresponding ILP problem. The results suggest that, if the associated weight to constraint (1) is larger than 1.2 ($w_1 > 1.2$), the following three grounding facts are likely to be true:

$$memberOf(Staff\_B, Dept\_B)$$
$$basedAt(Staff\_B, Office\_B)$$
$$hasSite(Dept\_B, Office\_B)$$

which indicates that the initial information implying that $Office\_B$ was assigned to $Dept\_A$ according to the design is not credible. A probable reason is the office was repurposed for the use of $Dept\_B$ after the completion and occupation of the building $bldg$.

Discussion

This paper demonstrates a feasible way to eliminate potential conflicts and acquire the 'most likely scenarios' by integrating multiple representations of building information. Leveraging the uncertain knowledge graph, the weights of facts and constraints indicate the credibility of such information and knowledge. Different weight assignments can lead to different MAP results. For example, if the weight assigned to constraint (1) in the case study is less than 1.2 ($w_1 < 1.2$), the most likely scenario becomes that $Staff\_B$ is employed by $Dept\_B$, but the workplace $Office\_B$ is allocated to $Dept\_A$, which looks weird. Therefore, the weight assignment is crucial and multiple factors must be considered. Generally, these factors include: (1) the data creation time, the older the information is, the less trustful it is; (2) the data source, automatically collected data is more trustful than manual input data; (3) the source where knowledge comes from, expert knowledge is more trustful than common sense. Besides, the assignment of the weights is important to maintaining the flexibility of space utilization information. For instance, if one staff can be employed by multiple departments simultaneously, the third constraint would become a soft fact with finite weight. In other words, the weight decides to what degree a logic formula is resisted. Any piece of information with finite weight can be rejected because the MAP inference pursues the globally optimal scenario depending on all assigned weights. Therefore, the weight assignment rules will be investigated further in future work. In addition to the space assignment, finance records and human resources employee directory, other datasets can be incorporated to enhance the reliability of the integrated information. These include the cost accounting information related to space usage and tenancy agreements.

BIM technology has drawn great attention since its first emergence. It is believed that BIM is not only a standard but a way of working with diverse building data. As a centralized project data depository, BIM integrates project information into a singular model, and stakeholders share their knowledge along with the development of buildings. However, even with the latest developments of cloud BIM and blockchain-based BIM, which accelerate a shift from the current centralized structure to a decentralized one, many industrial practitioners are sceptical about the disruptive changes that could exacerbate the existing data management processes, especially from the perspectives of data curation, data security, etc. Semantic web technology demonstrates its flexibility in representing information, which is preferable to some alternatives with rigid and complex hierarchical structures. Although using semantic web technologies can somehow compromise the integrity of asset information, the proposed approach can ensure information consistency by resolving conflicts among varied information to support core functions like space management.

According to the UK Government Data Quality Hub, information consistency is achieved when there is an absence of differences in multiple representations (Askham et al., 2013). Additionally, the information in a consistent manner enables the linking of information from multiple sources without verification, helping to gain efficiency when managing disparate information (Askham et al., 2013). Within the boundaries of this premise, this paper assumes that information consistency does not only limit to the values, but also includes the means of representation such as formats because of the heterogeneity of asset information. The data quality attribute of 'consistency' is linked to the quality dimension of 'accuracy' which corresponds to the right value in the correct representation (Askham et al., 2013). This implies that consistency is a prerequisite to other quality attributes such as 'accuracy' and 'timeliness'.

Conclusions

The fragmented and project-based nature of the building industry usually results in the duplication and redundancy of information from different domains, phases, and software tools. The potential conflicts within the collected information pose extra challenges when utilizing the information for supporting asset management processes. This paper proposes an uncertain knowledge graph-based method, which represents building information in the form of facts attached with associated weights, indicating their relative credibility. By conducting the Maximum A-Posterior query over the generated Markov Logic Network, the most likely and conflict-free information is acquired. The inference largely relies on the all given weights, which reflect the ascertainment of the credibility of facts in specific cases. The space management process is used to verify the feasibility of the proposed method. The semi-structured interviews with nine asset management professionals reveal that in practice, diverse information from disparate information systems is used to confirm space usage. The case study validates the proposed method can draw credible space information from dissimilar sources, which can reduce the time-consuming manual information verification processes.

Acknowledgments

The research leading to these results has been funded by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 860555.

References

Abdullah, S., Ali, H., and Sipan, I. (2013). Space management: Utilizing space usage. In The 21th IBIMA conference on Vision, pages 1232–1239.

Anand, S. K. and Kumar, S. (2022). Uncertainty analysis in ontology-based knowledge representation. New Generation Computing, 40(1):339–376.

Askham, N., Cook, D., Doyle, M., Fereday, H., Gibson, M., Landbeck, U., Lee, R., Maynard, C., Palmer, G., and Schwarzenbach, J. (2013). The six primary dimensions for data quality assessment. DAMA UK working group, pages 432–435.

Azizi, S., Nair, G., Rabiee, R., and Olofsson, T. (2020). Application of internet of things in academic buildings for space use efficiency using occupancy and booking data. Building and environment, 186:107355.

Becerik-Gerber, B., Jazizadeh, F., Li, N., and Calis, G. (2012). Application areas and data requirements for bim-enabled facilities management. Journal of construction engineering and management, 138(3):431–442.

Durdyev, S., Ashour, M., Connelly, S., and Mahdiyar, A. (2022). Barriers to the implementation of building information modelling (bim) for facility management. Journal of Building Engineering, 46:103736.

Fang, Z., Liu, Y., Lu, Q., Pitt, M., Hanna, S., and Tian, Z. (2022). Bim-integrated portfolio-based strategic asset data quality management. Automation in Construction, 134:104070.

He, S., Liu, K., Zhang, Y., Xu, L., and Zhao, J. (2014). Question answering over linked data using first-order logic. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), pages 1092–1103.

Hetherington, J. and West, M. (2020). The pathway towards an Information Management Framework-A 'commons' for Digital Built Britain. `https://doi.org/10.17863/CAM.52659`. Accessed: 2023-04-14.

ISO/IEC (2015). ISO/IEC 33001: Information technology-process assessment-concepts and terminology. `https://www.iso.org/standard/54175.html`. Accessed: 2023-04-14.

Ji, S., Pan, S., Cambria, E., Marttinen, P., and Philip, S. Y. (2021). A survey on knowledge graphs: Representation, acquisition, and applications. IEEE transactions on neural networks and learning systems, 33(2):494–514.

Li, X., Lyu, M., Wang, Z., Chen, C.-H., and Zheng, P. (2021). Exploiting knowledge graphs in industrial products and services: A survey of key aspects, challenges, and future perspectives. Computers in Industry, 129:103449.

Liu, L., Stumpf, A., Kim, S., and Zbinden, F. (1994). Capturing as-built project information for facility management. In Proceedings of the 1st Congress on Computing in Civil Engineering. Part 1 (of 2), pages 614–621. Publ by ASCE.

Pauwels, P., Costin, A., and Rasmussen, M. H. (2021). Knowledge graphs and linked data for the built environment. In Industry 4.0 for the Built Environment: Methodologies, Technologies and Skills, pages 157–183. Springer.

Pauwels, P., Zhang, S., and Lee, Y.-C. (2017). Semantic web technologies in aec industry: A literature overview. Automation in construction, 73:145–165.

Richardson, M. and Domingos, P. (2006). Markov logic networks. Machine learning, 62:107–136.

Rogers, C. (2002). Space Management in Higher Education: Report of the Findings of the Newcastle University Space Management Project, Jointly Funded by the HEFCE Good Management Practice Programme and the University (pp. 1-88). PhD thesis, Newcastle: University of Newcastle.

Spire, Z. (2022). University estates and postdigital higher education: Space, place, and being a university. Postdigital Science and Education, pages 1–14.

Valks, B., Arkesteijn, M. H., Den Heijer, A. C., and Putte, H. J. V. (2018). Smart campus tools–adding value to the university campus by measuring space use real-time. Journal of Corporate Real Estate, 20(2):103–116.

Zadeh, P. A., Wang, G., Cavka, H. B., Staub-French, S., and Pottinger, R. (2017). Information quality assessment for facility management. Advanced Engineering Informatics, 33:181–205.