2022 European Conference on Computing in Construction
Ixia, Rhodes, Greece
July 24-26, 2022

ENABLING DOWNSTREAM MACHINE-LEARNING OVER THE TEXTUAL INFORMATION CONTAINED IN BUILDING KNOWLEDGE GRAPHS

Mehrzad Shahinmoghadam[1], Ali Motamedi[1], and Mohammad Mostafa Soltani[2]
[1]École de technologie supérieure, Montreal, Canada
[2]Toronto, Canada

Abstract

The current practice of statistical learning from building information models mostly relies on the manual construction of table-oriented representation of the numeric data. This is while the lexical information contained in building information models can further reinforce the learning process by preserving the essential role of the semantic relationships. In this paper, application of one of the state-of-the-art knowledge graph embedding algorithms (RDF2Vec), yielded promising results with regards to an object clustering task. This paper contributes to the literature by shedding light on the potential of semantic embedding algorithms to facilitate downstream machine learning over building knowledge graphs.

Introduction

The volume of the research studies investigating the subject of *"Linked Building Data"* has been rapidly increasing over the past decade. To date, a significant portion of the research efforts were focused on the development of the representation schemas and knowledge models using formal ontologies and semantic web technologies. Such research efforts have resulted in a fair number of high-quality ontologies and rich schemas, some of which maintained and/or recommended by internationally known standardization organizations and consortiums (e.g., the ifcOWL ontology officially maintained by buildingSMART). The existing ontologies have proved effective in establishing the semantic relationships among the disconnected pieces of building data coming from heterogenous and disparate sources. By reusing, extending, or modifying, the existing ontologies and with the help of the right tools and technologies rich graphs of the built asset data can be created for various purposes (e.g., data integration, semantic query, knowledge sharing, etc.). Such graphs can be viewed as *"building knowledge graphs"*: graph-structured data representations that not only contain the data that describes the actual instances of the building entities (their properties and interrelationships), but also contain the background domain knowledge that reveals the inherent semantic relationships between the graph entities.

The existing body of knowledge on the subject of linked building data contains considerable research on the creation of semantically rich building graphs for various applications. The most common tools and technologies that were used in previous studies are among the ones that are officially recommended by W3C consortium as the semantic web standardized specifications, namely: OWL (web ontology language), RDF (Resource Description Framework), and SPARQL (a semantic query language for RDF triples). The ifcOWL ontology, an OWL-based representation of the Industry Foundation Classes (IFC) schema (*ifcOWL*, no date), has been commonly used to enable extra features (e.g., automated reasoning, highly expressive semantic query) for extending the general capabilities of Building Information Modeling (BIM), e.g., enhancing BIM interoperability with sensory operational data as previously reported in (Terkaj, Schneider and Pauwels, 2017; Zhong *et al.*, 2018; Shahinmoghadam and Motamedi, 2020).

Previous research has highlighted the automated reasoning power of the formal ontologies that can be used to automatically infer new facts about the building based on the predefined logical axioms (contained within the ontology). However, rule-based inference mechanisms suffer from known shortcomings (e.g., being hard to scale and low recall rate (Cui *et al.*, 2019)). Given such limitations, an in-depth analysis of the data contained in the building graphs can be performed through the use of statistical learning and pattern recognition techniques to discover new knowledge from graph-structured data. Yet, the latter has been disproportionately less investigated within the literature.

This is while the research on artificial intelligence has shown that knowledge graphs (i.e., formal descriptions of things and their interrelationships) can play a major role in creating hybrid intelligent systems that will significantly outperform traditional data processing systems. Such superiority of the knowledge graph-supported systems over traditional machine learning is due to their ability to enable knowledge-infused learning processes (Sikos, 2021). The ubiquity of network structures in the real world and the ability of the graph analysis/mining methods to uncover the interactions between the graph entities and understand the network behaviour in a systematic manner has been attracting an increasing attention in various research areas such as social network analysis, recommender systems, biology science, and linguistics, to name a few (Goyal and Ferrara, 2018).

In the context of the construction and building engineering research, Sacks et al. (Sacks, Girolami and Brilakis, 2020) recently pinpointed the significant potential of the formally structured data models for the representation of building models as labeled-property graph models. As mentioned by Sacks et al. (Sacks, Girolami and Brilakis, 2020), such graph models can facilitate machine learning tasks for semantic enrichment of building models (uncovering the implicit object relationships and properties within the model). The potential of graph-based representations for semantic enrichment of building information models has been most recently reported in (Wang, Sacks and Yeung, 2022). In another recent work, Abdelrahman et al. (Abdelrahman, Chong and Miller, 2020) presented a methodology for transforming labeled-property graphs obtained from IFC-based BIM models and temporal sensor data into vector spaces. In that study, the researchers used two types of graph embedding algorithms for this purpose and promising results were reported accordingly. The existence of such studies, although still extremely rare within the literature, indicate to the fact that the research community has started to explore the potential of graph data science for building analytics and BIM research.

This paper reports on initial results obtained from an ongoing research study that seeks to examine the potential usefulness of knowledge graph embedding algorithms to facilitate downstream statistical analysis of the data contained within RDF-based knowledge graphs, by learning vector representations of the graph entities. For the current study, we used RDF2Vec (Ristoski *et al.*, 2019): an unsupervised, task-agnostic algorithm that is capable of representing the entities of a given RDF graph as a vector of numerical values. Using an existing RDF dataset of a BIM model, a case study was conducted to examine the quality of the derived embeddings in an object clustering scenario. The results of the case study indicate that the derived embeddings had sufficient discriminative representation power to be used for obtaining fair clusters of beam and wall elements within the examined knowledge graph. To the best of our knowledge, to date, no other study has empirically examined the potential of the knowledge graph embedding algorithms for RDF-based graphs that represent BIM models.

Background

To establish the key terminology used throughout this paper, definition of preliminary concepts and a short overview of the essential background information is provided as follows.

Knowledge graphs

According to a recent survey (Hogan *et al.*, 2021), various proposals have been emerging for the definition of a knowledge graph. Pauwels et al. (Pauwels, Costin and Rasmussen, 2022) have most recently conducted a thorough investigation of knowledge graphs and linked data in the context of the built environment. The interested reader is referred to their work to gain a profound understanding of the basics of knowledge graphs, the enabler technologies, and their usefulness for the AEC and smart building applications.

For the purpose of the current study, we borrowed the definitions provided in (Sikos, 2021) and (Hogan *et al.*, 2021): A knowledge graph is a graph of data, often stored in graph databases or triplestores, whose nodes contain information about real-world entities and whose edges describe the interrelationships between those entities.

Based on the above definition, an RDF representation of a BIM model which will be supported with ontologies (e.g., ifcOWL) acting as shared agreed-upon vocabularies and formal knowledge models, can be viewed as a *"knowledge graph"* which describes a specific building and its properties (e.g., object geometry, system design specifications, material properties, etc.). However, it should be noted that compared to large-scale general knowledge graphs such as DBpedia (Auer *et al.*, 2007), RDF graphs describing individual or districts of buildings will be categorized as small to medium scaled knowledge graphs.

Knowledge graph embedding

With reference to the definitions presented in [6] and [11], a graph embedding is a mapping of the knowledge graph components (including the nodes and the relations) to a low-dimensional feature vector, while preserving the connection strengths and the inherent structure of the original graph.

According to the above definition and in the context of the current study, the graph embeddings can be deemed as dense, lower-dimensional representations of the building entities that were originally described by a set of RDF triples. Hence, no matter if a given entity of the building graph has been described by words (described by triples consisting of string values) or by numerical data (described by triples consisting of numerical values), that entity can be quantitatively described with the help of a purely numeric representation in a vector space. This way, downstream quantitative analysis of the building graph data will be significantly simplified.

Graph embedding algorithms

Graph embedding algorithms can be defined as computational methods capable of learning efficient feature representations that can be used for mapping the graph components into low-dimensional continuous vector spaces (Wang *et al.*, 2017). By using such algorithms, the numerical representations of the graph data in the form of the entity embeddings can be effectively used for various downstream tasks such as clustering, node classification, visualization, and link prediction (Goyal and Ferrara, 2018).

In this study, the RDF2Vec algorithm (Ristoski *et al.*, 2019) was used to derive the embeddings for the building graph components in a case study. An overview of the RDF2Vec approach can be given as follows: the algorithm adapts neural language models to derive numerical representations of the entities in RDF graphs. To this end, the algorithm first transforms the graph data into sequences of entities, which can be deemed as word sentences. Then, taking the assumption that closer words in a sequence are statistically more dependent, it trains neural language models to represent each entity in the RDF graph as a vector of numerical values in a latent feature space [9].

Potential benefits and applications

Knowledge graphs and their embeddings can bring valuable benefits to data mining and knowledge extraction from building lifecycle data. This section gives an overview of such potential benefits and provides some application examples that will be of interest to the practice.

Since knowledge graphs can enrich building data with *"context"* (Pauwels, Costin and Rasmussen, 2022), they can be used to reinforce the statistical learning processes by taking advantage of the semantic descriptions that can be found for each entity of the data within the knowledge graphs.

Given the known challenges of representing the building's geometry and operational data (time-series observations) using the linked data technologies such as RDF (Pauwels, Costin and Rasmussen, 2022), building knowledge graphs mostly consist of textual information. In this light, application of knowledge graph embedding algorithms offers a vast opportunity to enable an in-depth analysis of the content of the knowledge graphs, at both lexical and semantic levels, i.e., knowledge graph embeddings facilitate downstream machine learning over the textual information that is contained in the building knowledge graph.

As the first step, knowledge graphs can be used to augment the building datasets with background knowledge (data contextualization). As the next step, by learning the semantic embeddings of the entities of interest through the use of the embedding algorithms, dense representations of the contextualized data can be obtained and represented in a continuous vector space. Finally, the state-of-the-art data mining and machine learning methods can be applied over the dense and vectorized representations of the knowledge graph content to extract new knowledge from the contextualized data.

As an immediate application, knowledge graph embedding algorithms can be applied to assist the schema matching task. In particular, various machine learning algorithms can be applied over the learnt semantic embeddings of the various schemas used in AEC/FM domains to find the semantic correspondences of the entities between two or more schemas. Such application will be of high value to the integration of the disparate sources of the building data (e.g., BIM and sensor data integration (Shahinmoghadam and Motamedi, 2020)). As another important application, knowledge graph embeddings can be used to facilitate the semantic enrichment of the building information models. For example, the semantic embeddings of the building entities can be used to augment the content of an existing knowledge graph with semantic tags (e.g., annotating an ifcOWL-based representation with the tags from the Brick schema (Fierro *et al.*, 2019)). Similarly, knowledge graph embeddings can assist in the tasks where context plays an essential role. An interesting potential scenario in this regard is to establish the semantic correspondence between the 3D and scheduling information in the context of 4D-BIM. Finally, knowledge graph embeddings can make direct contributions to the development of Question-Answering systems over building knowledge graphs. In this regard, the recent research in the computer science domain has shown that semantic embeddings of the knowledge graphs can effectively enhance the results of information retrieval and query relaxation tasks (Mai *et al.*, 2020; Chen *et al.*, 2021).

Experimental demonstration

Approach

Figure 1 illustrates the main steps of the approach that has been adopted in this study to examine the usefulness of the knowledge graph embedding algorithms in the context of the construction and building engineering research. Although the body of the literature on graph analysis is abundant with various approaches and algorithms for graph embedding purposes (the interested reader is referred to the survey reported in (Goyal and Ferrara, 2018)), RDF2Vec is one of the few methods that has proved effective for meeting the specific characteristics of the graphs that are represented using the RDF technology. Hence, we were motivated to use this algorithm to meet the objective of the present study, i.e., creating embeddings of the entities contained in RDF-formatted BIM data. For the sake of quantitative examinations, in this study we used a Python implementation of RDF2Vec which is publicly available at [13].

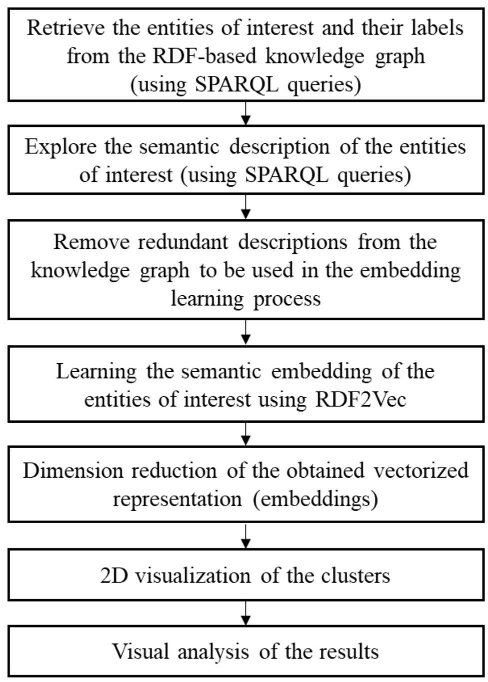

Figure 1: The main steps taken in this study to conduct the experimental evaluations

To create a readily usable dataset, we performed the pre-processing of the knowledge graph that was used in this case study. To this end, we formulated different sets of SPARQL queries to perform an exploratory analysis of the content of the knowledge graph. More details on the description of the data used in this work is given in the next section.

Data

In this study we used an existing RDF-formatted knowledge graph which is publicly available at an online repository maintained by a research lab at Ghent University, Belgium [14]. The used model is an RDF representation of an IFC BIM model which has been created with the purpose of representing the structural elements of a general building. Figure 2 shows the IFC-based visualization of the BIM model used in this work. Both IFC version (.ifc format) and RDF version (.ttl format) of the used model are publicly accessible at the mentioned repository under the title "02_BIMcollab_Example_STR". The ifcOWL ontology (IFC2X3_TC1 version) is the main vocabulary (schema) that has been used to describe the graph entities in this RDF dataset. For the manipulation of the used RDF triples, we used RDFLib [15] which is a powerful and widely-used Python library for working with RDF data.

Figure 2. IFC-based visualization of the building RDF dataset used for the case study

For the purpose of examining the effectiveness of the adopted algorithm, we decided on learning the vectorized representation of two specific structural components, i.e., beams and walls. The rationale behind this decision is to evaluate the discriminative representation power of the semantic embeddings that were learnt merely from the textual descriptions of each element object (i.e., when geometric descriptions as the fundamental characteristics of beams and walls were absent from the building information model). To extract the entities of interest, two sets of SPARQL queries were run over the data set to identify all object instances of types "ifcowl:IfcBeam" and "ifcowl:IfcWallStandardCase". After running the queries, 91 beam and 114 wall elements were found in the dataset. The results of the query were saved in the ".csv" format to be used as input to the algorithm as identifiers of the entities for which the embeddings should be computed. Moreover, other sets of SPARQL queries were formulated to explore the sematic descriptions that could be found for each entity of interest within the original knowledge graph. The results of the queries were also used to identify redundant descriptions that could be removed before starting the embedding learning process. As a result, two sets of predicates describing the owner history and object global ID were skipped from the original knowledge graph as they were of less informative value.

Evaluation

To empirically examine the usefulness of the selected approach, the graph embeddings derived after applying the RDF2Vec algorithm over the RDF dataset were used as input for a clustering algorithm. To this end, two types of structural building components, namely "beams" and "walls", were considered to be included in the object clustering scenario. In order to be able to visually analyze the discriminative power of the derived embeddings, the dimensionality of the embeddings was reduced in a way to be able to create 2D data representation plots for the transformed entities (i.e., beams and walls). For this purpose, we used t-SNE algorithm (t-distributed Stochastic Neighbour Embedding) [16], which is a powerful tool for creating intuitive visualizations of high-dimensional data. A Python implementation of this algorithm which is included in the Scikit-learn data science package was used in this study.

Results

Prior to running the embedding algorithm, the RDF triples mentioning the type of the entities (i.e., triples with "rdf:type" as their predicate) were skipped to be included within the learning process. This way, the trained model was kept blind from the actual entity type labels (beam and wall type labels).

Subsequent to running the embedding algorithm over the building dataset, the main output which is the embeddings of the identified entities were saved to be used in the object clustering step. The derived embeddings are in fact the numerical representations of the beams and wall nodes with accordance to the feature vectors that were computed through the unsupervised training of the neural models based on the descriptions provided within the building's RDF graph. Finally, the dimensionalities of the obtained beam and wall entities were reduced using the t-SNE technique.

Figure 3 and Figure 4 show the 2D plots of the obtained beam and wall embeddings, respectively. In order to be able to more clearly inspect the effectiveness of the output for the object type clustering purpose, all embeddings were plotted in Figure 4 (the object unique IDs were removed for better readability).

As it can be seen from Figure 3 and Figure 4, the beam embedding representations are more dense at the lower left region of the plot, while the wall embedding representations are more closely packed at the upper right region of the plot.

Although looking at the beam and wall embedding plots individually does not clearly reflect the data density distributions for the beam and wall objects, by looking at Figure 5, one can more promptly notice the difference between beam and wall density distributions on the plot.

When looking at the obtained embedding plots, it must be noted that these representations were created after reducing the dimensionality of the derived embeddings, which was done for the sake of intuitive visual inspection of the results. However, the discriminative power of the derived numeric representations (embeddings) will be increased when higher dimensions of the learnt feature vectors are included within the computations for downstream analysis of the entity embeddings.

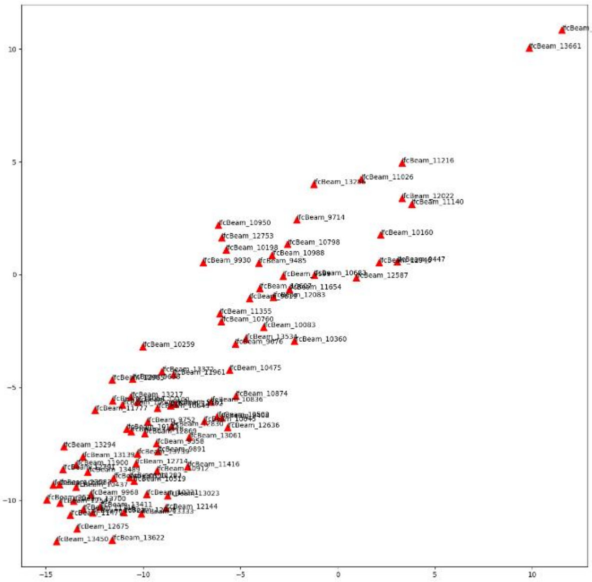

Figure 3. 2D visualization of the "beam" embeddings

To perform a more in-depth analysis of the results, the following steps were taken for this case study: using the IFC-formatted model which corresponds to the RDF dataset that was used in this work, we started to manually inspect the wall and beam elements of the building in an IFC viewer tool. During this inspection process, it was found that there exist a number of walls that are visually distinctive from the rest of the wall cases contained in the BIM model. To give an example, between the two wall elements highlighted in Figure 6, the wall that is marked with number "1" (which seems to be a foundation beam rather than a standard wall case), resembles more similarities to a beam element than to a wall in comparison with the other highlighted wall (the one marked with number "2" in the figure).

To locate the embedding representation of the slender-shaped wall (marked with number "1" in Figure 6) on the wall embeddings plot (Figure 4), we tracked its corresponding RDF instance by matching the GUID of the wall element that was extracted from the original IFC model. After matching the extracted GUID with the RDF descriptions contained in the dataset, we were able to locate the wall embedding of interest which is marked with the red arrow in Figure 4.

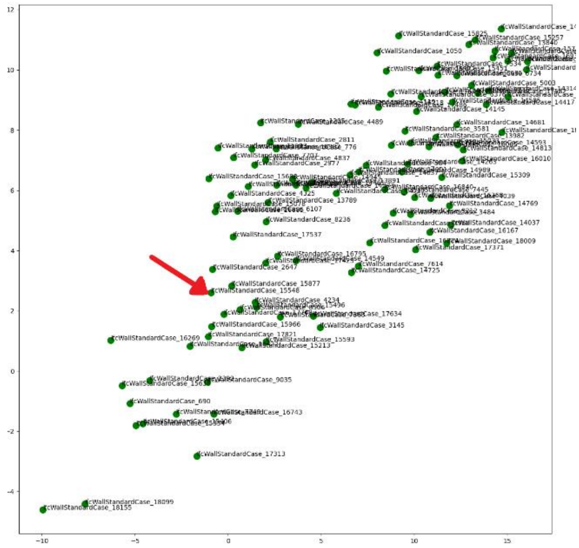

Figure 4. 2D visualization of the "wall" embeddings

Now, a closer look at Figure 4 and Figure 5 reveals that the embedding location of the tracked wall is farther from the dense area of wall embedding representations and closer to the region that beam embeddings express more density on the plot. Evidently, this observation is in accordance with our initial intuitive interpretation that was formed on the basis of the visual inspection of the wall elements. However, considering the fact that the obtained embedding results were not derived based on the detailed geometry data (the numerical values of the element dimensions were not included within the used RDF dataset at the first place), the used algorithm has been effectively capable of detecting the hidden inherent similarities between the slender walls and beam elements in general.

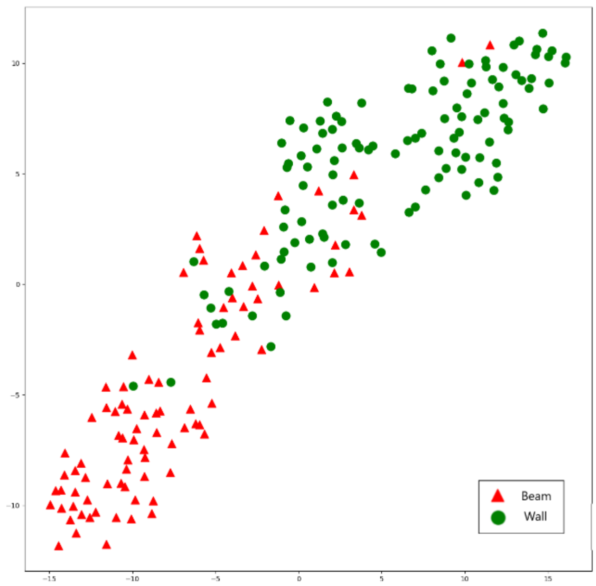

Figure 5. Derived clusters for beam and wall entities

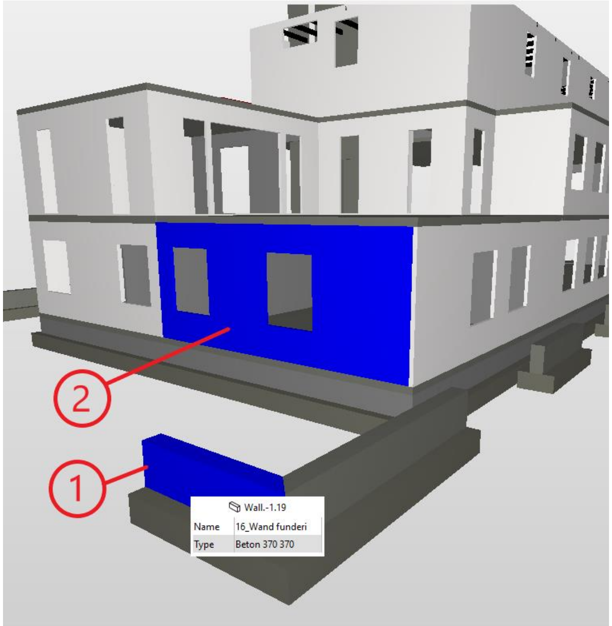

Figure 6. Using IFC-based representations for visual inspection of the building elements

To conclude the discussion of the results, regardless of the small size of the dataset used in this case study, the results appeared to be promising with regards to the derivation of the numerical representations of the linked building graph data (RDF graphs) in vector spaces. Although the derived embeddings are the results of an unsupervised learning algorithm and the interpretation of the exact mechanism of such learning processes remains unclear to our understanding (similar to the majority of the state-of-the-art machine learning algorithms) more profound analyses of the graph data might provide some clues to explain why some of the embeddings were slightly scattered far from the dense regions in beam and wall embedding plots. However, the observed scatteredness can be explained by the fact that the size of the RDF dataset used in this study has been relatively small. In this regard, not having access to a large and high-quality dataset of graph-structured building data in the form of RDF-based knowledge graphs has been one of the major challenges that we faced in this research study. Not only the scarcity of publicly available BIM models is one of the main drivers of the aforementioned challenge, but also the limited number and practical capabilities of the existing tools that can be used to transform conventional BIM models into semantically rich graph-based representations is adding to the severity of the challenge. Hence, since creation of a large and high-quality RDF dataset will require considerable time, effort, and resources, further investigations with regards to the effectiveness of the used algorithm, as well as other existing graph embedding algorithms (as potential alternatives) will be beyond the scope of the current study.

Finally, it is important to note that for the case study reported here, no parameter tuning was performed in advance to the analysis of the results. Hence, optimization of those parameters that have an influence on the output quality (e.g., graph walking strategy, number of graph walks, etc.) remains to be investigated through the future research using larger datasets with higher qualities. Moreover, for the future works, the usefulness of the embedding results should be tested in more complex scenarios such as multi-object classification or clustering, and link prediction

Conclusion

Statistical learning from linked building data graphs has been rarely investigated within the existing body of knowledge. The study reported in this paper sets out to answer the question "How the background knowledge embedded within the ontology-based building data representations (building knowledge graphs), can be exploited to reinforce the downstream statistical analysis of the graph-represented building data?"

In this paper, we examined a graph embedding algorithm that has been originally proposed to facilitate derivation of vector representations for RDF graph components in continuous vector spaces. To test the quality of the results, a case study was conducted to use the derived embeddings for the purpose of creating clusters of two building structural elements. Despite the small size of the used dataset, the obtained results were promising.

This work makes original contributions to the literature by elaborating on the potential benefits of applying semantic embedding algorithms over building knowledge graphs and its implications to the AEC and smart building domains. To allow other researchers to reproduce the results and/or explore the approach in their research, the codes used for the exploratory analysis of the knowledge graph, dataset creation, and learning the semantic embeddings of the graph entities will be made available at (https://github.com/ccps-lab).

The future research will focus on the development of large and high-quality datasets to test and compare the effectiveness of the state-of-the-art knowledge graph embedding algorithms in more complex scenarios such as automated semantic tagging of the building graph data, automated quality checking of the building information models, and Question-Answering over building knowledge graphs.

References

Abdelrahman, M., Chong, A. and Miller, C. (2020) 'Build2Vec: Building Representation in Vector Space', arXiv:2007.00740 [cs] [Preprint]. Available at: http://simaud.org/2020/proceedings/102.pdf (Accessed: 18 May 2021).

Auer, S., Bizer, C., Kobilarov, G., Lehmann, J., Cyganiak, R. and Ives, Z. (2007) 'Dbpedia: A nucleus for a web of open data', in The semantic web. Springer, pp. 722–735.

Chen, J., Hu, P., Jimenez-Ruiz, E., Holter, O.M., Antonyrajah, D. and Horrocks, I. (2021) 'OWL2Vec*:

embedding of OWL ontologies', Machine Learning, 110(7), pp. 1813–1845. doi:10.1007/s10994-021-05997-6.

Cui, W., Xiao, Y., Wang, H., Song, Y., Hwang, S. and Wang, W. (2019) 'KBQA: learning question answering over QA corpora and knowledge bases', arXiv preprint arXiv:1903.02419 [Preprint].

Fierro, G., Koh, J., Agarwal, Y., Gupta, R.K. and Culler, D.E. (2019) 'Beyond a House of Sticks: Formalizing Metadata Tags with Brick', in Proceedings of the 6th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation. New York, NY, USA: Association for Computing Machinery (BuildSys '19), pp. 125–134. doi:10.1145/3360322.3360862.

Goyal, P. and Ferrara, E. (2018) 'Graph embedding techniques, applications, and performance: A survey', Knowledge-Based Systems, 151, pp. 78–94. doi:10.1016/j.knosys.2018.03.022.

Hogan, A., Blomqvist, E., Cochez, M., D'amato, C., Melo, G.D., Gutierrez, C., Kirrane, S., Gayo, J.E.L., Navigli, R., Neumaier, S., Ngomo, A.-C.N., Polleres, A., Rashid, S.M., Rula, A., Schmelzeisen, L., Sequeda, J., Staab, S. and Zimmermann, A. (2021) 'Knowledge Graphs', ACM Computing Surveys, 54(4), p. 71:1-71:37. doi:10.1145/3447772.

ifcOWL (no date) buildingSMART Technical. Available at: https://technical.buildingsmart.org/standards/ifc/ifc-formats/ifcowl/ (Accessed: 4 August 2021).

Mai, G., Yan, B., Janowicz, K. and Zhu, R. (2020) 'Relaxing Unanswerable Geographic Questions Using A Spatially Explicit Knowledge Graph Embedding Model', in Kyriakidis, P., Hadjimitsis, D., Skarlatos, D., and Mansourian, A. (eds) Geospatial Technologies for Local and Regional Development. Cham: Springer International Publishing (Lecture Notes in Geoinformation and Cartography), pp. 21–39. doi:10.1007/978-3-030-14745-7_2.

Pauwels, P., Costin, A. and Rasmussen, M.H. (2022) 'Knowledge Graphs and Linked Data for the Built Environment', in Bolpagni, M., Gavina, R., and Ribeiro, D. (eds) Industry 4.0 for the Built Environment: Methodologies, Technologies and Skills. Cham: Springer International Publishing (Structural Integrity), pp. 157–183. doi:10.1007/978-3-030-82430-3_7.

Ristoski, P., Rosati, J., Di Noia, T., De Leone, R. and Paulheim, H. (2019) 'RDF2Vec: RDF graph embeddings and their applications', Semantic Web, 10(4), pp. 721–752. doi:10.3233/SW-180317.

Sacks, R., Girolami, M. and Brilakis, I. (2020) 'Building Information Modelling, Artificial Intelligence and Construction Tech', Developments in the Built Environment, 4, p. 100011. doi:10.1016/j.dibe.2020.100011.

Shahinmoghadam, M. and Motamedi, A. (2020) 'An Ontology-Based Mediation Framework for Integrating Federated Sources of BIM and IoT Data', in Proceedings of the 18th International Conference on Computing in Civil and Building Engineering. International Conference on Computing in Civil and Building Engineering, Springer, Cham, pp. 907–923. doi:10.1007/978-3-030-51295-8_63.

Sikos, L.F. (2021) Provenance in Data Science: From Data Models to Context-Aware Knowledge Graphs. Springer Nature.

Terkaj, W., Schneider, G.F. and Pauwels, P. (2017) 'Reusing domain ontologies in linked building data : the case of building automation and control', in Proceedings of the Joint Ontology Workshops 2017 Episode 3: The Tyrolean Autumn of Ontology. 8th International Workshop on Formal Ontologies meet Industry. Available at: http://hdl.handle.net/1854/LU-8578705 (Accessed: 15 November 2019).

Wang, Q., Mao, Z., Wang, B. and Guo, L. (2017) 'Knowledge Graph Embedding: A Survey of Approaches and Applications', IEEE Transactions on Knowledge and Data Engineering, 29(12), pp. 2724–2743. doi:10.1109/TKDE.2017.2754499.

Wang, Z., Sacks, R. and Yeung, T. (2022) 'Exploring graph neural networks for semantic enrichment: Room type classification', Automation in Construction, 134, p. 104039. doi:10.1016/j.autcon.2021.104039.

Zhong, B., Gan, C., Luo, H. and Xing, X. (2018) 'Ontology-based framework for building environmental monitoring and compliance checking under BIM environment', Building and Environment, 141, pp. 127–142. doi:10.1016/j.buildenv.2018.05.046.