Project Work Breakdown Structure Similarity Estimation Using Semantic and Structural Similarity Measures

Navid Torkanfar[a] and Ehsan Rezazadeh Azar[a]

[a] Department of Civil Engineering, Lakehead University, Canada
E-mail: ntorkanf@lakeheadu.ca, eazar@lakeheadu.ca

Abstract –

Reusing of the past information and lessons learned helps practitioners in better management of various aspects of construction projects, such as cost estimation, planning, contracting, and design. Measuring the similarity of construction projects improves the efficiency of the existing information systems in retrieval of relevant cases. It was hypothesized that the Work Breakdown Structure (WBS) of projects contains the necessary information to measure the semantic similarity of construction projects; therefore, WBS can be used as a potential representative of the projects. In this research project, a novel method is proposed to assess the semantic similarity of projects by application of natural language processing techniques. In this method, a new project is compared with the documented as-built projects based on their WBS similarity. This method is implemented using two metrics: (1) node similarity that compares the semantics of all nodes in two WBSs; (2) structural similarity which compares the topology of the work breakdown structures. The proposed system calculates a similarity score between 0 and 1 for each metric and the combination of these two scores provides the final similarity score between a pair of WBSs, thus it could rank the similarity of the documented cases to the new project based on their final scores. Experimental results indicated that the structural similarity produced about 15 percent higher degree of retrieval precision than the node similarity.

Keywords –

Project similarity; Construction project; Work breakdown structure; Natural language processing; Knowledge management

1 Introduction

Effective reuse of gained knowledge from past projects helps managers to complete projects in a timely and economical manner, and with a higher quality [1]. In the construction industry, several methods such as knowledge management techniques and case-based reasoning (CBR) have been applied to reuse past information and experiences. A major step in the knowledge retrieval systems is to index stored data based on several attributes in order to find the most relevant document(s).

Knowledge management systems are IT-based systems that aim to improve the organizational process of knowledge creation, storage/retrieval, transfer, and application [2]. The process of knowledge retrieval is difficult and can result in irrelevant documents [3]. A method was proposed to retrieve relevant information in construction projects by keywords, such as project type and title, through a Google-like function [4].

A CBR system can offer a solution for new problems by recalling a similar past situation(s) [5]. CBR methods have been investigated in different areas of construction research, including cost estimation [6,7], safety [8], and planning [9,10]. In CBR systems, the similarity of cases is compared based on predefined nominal or numerical attributes. Finding the appropriate attributes and assigning weights are among the main challenges in these systems.

A quantitative similarity measurement among construction projects provides a comprehensive metric that can improve the process of retrieving related information. Current practices in similarity measurement of construction projects are limited. These methods distinguish construction projects based on generic factors, such as project size and location or specific user defined attributes, such as number of floors, structural system type or keywords, rather than the entire scope of the project The aim of this research is to fill this gap by proposing a method to compare projects using their WBSs.

WBS is one of the main outcomes of the processes in the scope management of projects [11]. The WBS is used in different project management areas such as time and cost management [11]. The proposed system is the first of the kind attempt to use WBS to measure the similarity of the project scopes. The metrics for this assessment were developed using NLP techniques. NLP was employed to semantically compare the tasks and services within the WBSs.

1.1 Work breakdown structure (WBS)

WBS is s a hierarchical decomposition of the tasks and services to be carried out by the project team to achieve the project objectives [11]. WBS hierarchy begins with the project name at its highest level. The tasks and services in each level of the hierarchy are subdivided into smaller tasks that must be accomplished to satisfy the higher-level packages (samples in Figure 1). This decomposition continues until the tasks cannot be broken down or are no longer meaningful.

1.2 Semantic similarity measurements

The aim of the studies in NLP area is to enable computers to understand natural language text and speech used in human communications [12]. Measuring the similarity between words and sentences is an NLP technique that has been used in different areas, such as text classification, document clustering, and text summarization [13]. Corpus-based [14] and knowledge-based systems are utilized to semantically measure the similarity of words based on their definitions.

In the knowledge-based systems, semantic similarities are measured based on semantic relations among concepts [15]. These relations are embedded in the semantic networks. WordNet, a lexical database of English, is one of the most popular semantic networks [16]. The hierarchical structure of WordNet contains various semantic relations, such as synonymy, autonomy, hyponymy, and membership. These relations are used to calculate semantic similarities [17].

Wu and Palmer method (WUP) is a procedure to measure the semantic similarities between English words based on the WordNet [18]. In this method, the similarity is measured based on the position of concepts c_1 and c_2 and the position of their lowest common subsumer $lso\ (c_1, c_2)$ in the WordNet hierarchy. For example, Equation 1 calculates the similarity of concepts c_1 and c_2 , where the $len(c_1, c_2)$ is the length of the shortest path between concepts c_1 and c_2, and the $depth$ is the length of the path that connects each concept to the root element [18].

$$sim_{WUP}(c_1, c_2) = \frac{2 * depth(lso\ (c_1, c_2))}{len(c_1, c_2) + 2 * depth(lso\ (c_1, c_2))} \quad (1)$$

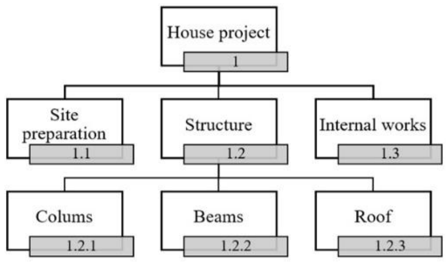
2 Method

The proposed method quantifies the similarity between two construction projects by two metrics driven from their WBSs: node similarity and structural similarity. These metrics are calculated based on semantic comparisons of tasks within the two compared WBSs. These semantic comparisons include three measurements: the semantic comparison of nodes, parents of the nodes, and siblings of the nodes.

2.1 Semantic measurements

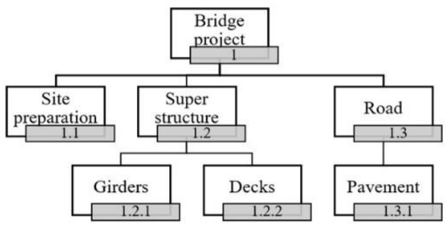
The tasks in a WBS are composed of several words.

Task Code	Task Name
1	House project
1.1	Site preparation
1.2	Structure
1.2.1	Columns
1.2.2	Beams
1.2.3	Roof
1.3	Internal works



(a)

Task Code	Task Name
1	Bridge project
1.1	Site preparation
1.2	Super structure
1.2.1	Girders
1.2.2	Decks
1.3	Road
1.3.1	Pavement



(b)

Figure 1. Two simplified work breakdown structures: (a) a House project (b) a Bridge project

The similarity between tasks can be calculated by averaging semantic similarity of their words. This study utilizes Equation 2 [19] to calculate the similarity between $task_i$ and $task_j$. In this equation, the method finds the most similar word for each word (w) in $task_i$ from $task_j$ ($maxSim\left(w, task_j\right)_{wup}$) based on Wu and Palmer algorithm [18].

$$sim_{semantic}\left(task_i, task_j\right) = \frac{1}{2}\left(\frac{\sum_{w\in\{task_i\}}\left(maxSim\left(w, task_j\right)_{wup}\right)}{\sum_{w\in\{task_i\}} 1} + \frac{\sum_{w\in\{task_j\}}\left(maxSim\left(w, task_i\right)_{wup}\right)}{\sum_{w\in\{task_j\}} 1}\right) \tag{2}$$

2.2 Comparison of nodes

A WBS is made of a hierarchy of nodes containing all the tasks required to complete a project. As shown in Figure 1, each node contains a task name and a task code. The first step in measuring the similarity between two WBSs is to semantically compare the task names. In this study, WBS_N^L represents an entire WBS with N nodes and L levels of hierarchy. Elements of the matrix in Equation 3 represent the semantic similarities between nodes of $WBS_{N_1}^{L_1}$ and $WBS_{N_2}^{L_2}$. In this equation $sim_{semantic}\left(n_i\ , m_j\right)$ equals to the semantic similarity between tasks of nodes n_i and m_j. N_1 and N_2 represent the list of nodes that each WBS contains (N_1: $(n_1, n_2, \ldots, n_N)$ and N_2: $(m_1, m_2, \ldots, m_M)$).

$$sim_{nodes}\left(WBS_{N_1}^{L_1}, WBS_{N_2}^{L_2}\right) = \begin{bmatrix} sim_{semantic}(n_1\ , m_1) & \cdots & sim_{semantic}\left(n_1\ , m_j\right) \\ \vdots & \ddots & \vdots \\ sim_{semantic}(n_i\ , m_1) & \cdots & sim_{semantic}(n_N\ , m_M) \end{bmatrix} \tag{3}$$

2.3 Comparison of node parents

In a WBS, each node is subdivided from an upper-level node which is called the parent of that node. Parent comparison between two nodes measures the semantic similarity of their parents. As shown in Figure 2, each node has a sequence of parents, and the first two parents with a similarity less than a specified threshold are defined as the least similar parents (LSP). The parent similarity, which is calculated using Equation 4, measures the arithmetic mean of semantic similarity of parents which are placed between the intended node and its LSP.

$$sim_{parents}\left(n_i\ , m_j\right) = \frac{\sum_{i=1}^{L_{LSP}-L_n}\left(L_{LSP} - L_{n_i} - (i-1)\right) \times sim_{semantic}(ith\ parents)}{\sum_{i=1}^{L_{LSP}-L_{n_i}}(i)} \tag{4}$$

The matrix in Equation 5 contains the parent similarities between nodes of $WBS_{N_1}^{L_1}$ and $WBS_{N_2}^{L_2}$.

$$sim_{parents}\left(WBS_{N_1}^{L_1}, WBS_{N_2}^{L_2}\right) = \begin{bmatrix} sim_{parents}(n_1\ , m_1) & \cdots & sim_{parents}\left(n_1\ , m_j\right) \\ \vdots & \ddots & \vdots \\ sim_{parents}(n_i\ , m_1) & \cdots & sim_{parents}(n_N\ , m_M) \end{bmatrix} \tag{5}$$

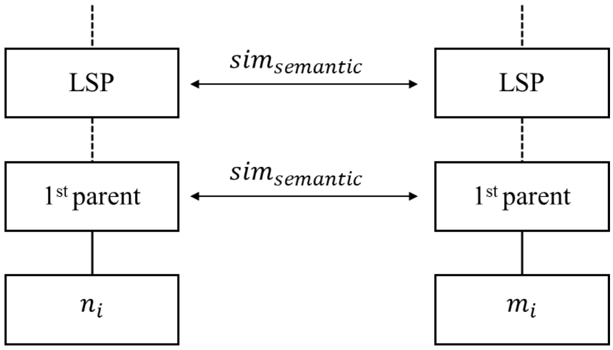


Figure 2. Parent similarity between nodes $\boldsymbol{n_i}$ and $\boldsymbol{m_j}$

2.4 Comparison of node siblings

In a WBS hierarchy, the nodes subdivided from a common parent are called siblings. The sibling similarity between nodes n_i and m_j is calculated by Equation 6 and is the fraction of the number of siblings that were actually matched by the total number of siblings. As shown in Figure 3, any pairs of siblings with a semantic similarity more than a defined threshold are considered matched. The matrix in Equation 7 contains the sibling similarities between nodes of $WBS_{N_1}^{L_1}$ and $WBS_{N_2}^{L_2}$.

$$sim_{siblings}\left(n_i, m_j\right) = \frac{\left|matched_{siblings}\left(n_i, m_j\right)\ \right|}{\left|siblings_{n_i}\right| + \left|siblings_{m_j}\right|} \tag{6}$$

$$sim_{siblings}\left(WBS_{N_1}^{L_1}, WBS_{N_2}^{L_2}\right) = \begin{bmatrix} sim_{siblings}(n_1\ , m_1) & \cdots & sim_{siblings}(n_1\ , m_M) \\ \vdots & \ddots & \vdots \\ sim_{siblings}(n_i\ , m_1) & \cdots & sim_{siblings}(n_N\ , m_M) \end{bmatrix} \tag{7}$$

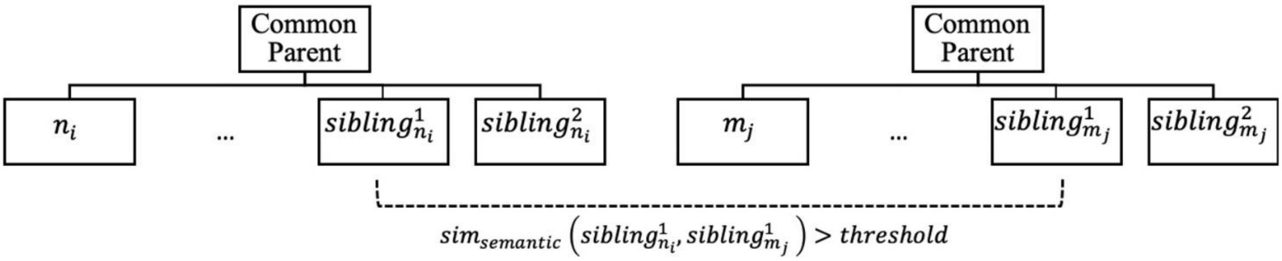


Figure 3. Sibling similarity between nodes $\boldsymbol{n_i}$ and $\boldsymbol{m_j}$

2.5 Mapping of nodes

The average of node, parent, and sibling similarities results in the average similarities (see Equation 8). The matrix in Equation 9 contains the average similarities between nodes of $WBS^{L_1}_{N_1}, WBS^{L_2}_{N_2}$. In this matrix, for instance, the first row includes the similarities between node n_1 from $WBS^{L_1}_{N_1}$, and nodes m_1 to m_M from $WBS^{L_2}_{N_2}$. The node n_1 will be mapped to a node from the second WBS with the highest average similarity. The same procedure will map all the nodes between two WBSs.

Mapped nodes in Equation 10 contain tuples of nodes which are mapped together $\left(n_i, m_j, sim_{average}\right)$ with their average similarity score (i.e. $sim_{average}$). Mapped nodes and $sim_{average}$ will be used to calculate the similarity scores between two WBSs, as they are utilizied to calculate node similarity and structural similarity scores ($n_i \in N_1$ and $m_j \in N_2$).

$$sim_{average} = \frac{sim_{node} + sim_{parents} + sim_{siblings}}{3} \tag{8}$$

$$sim_{average}\left(WBS^{L_1}_{N_1}, WBS^{L_2}_{N_2}\right) = \begin{bmatrix} sim_{average}(n_1, m_1) & \cdots & sim_{average}(n_1, m_j) \\ \vdots & \ddots & \vdots \\ sim_{average}(n_i, m_1) & \cdots & sim_{average}(n_N, m_M) \end{bmatrix} \tag{9}$$

$$mapped\ nodes = \left\{\left(n_i, m_j, sim_{average}\right)\right\} \tag{10}$$

2.6 Node similarity score

The node similarity score is the arithmetic mean of $sim_{average}$ of the mapped nodes. The node similarity, a score between 0 and 1, is calculated using Equation 11. In this equation, the sum of $sim_{average}$ is divided by the total number of nodes between two work breakdown structures.

$$Node\ similarity\left(WBS^{L_1}_{N_1}, WBS^{L_2}_{N_2}\right) = \frac{2 \times \sum_{(n_i, m_j) \in mapped\ nodes}\left(sim_{avg}\left(n_i, m_j\right)\right)}{|N_1| + |N_2|} \tag{11}$$

2.7 Structural similarity score

In the comparison of two WBSs, in addition to the semantic similarity of nodes within the WBS, the structure of hierarchies affects the similarity score as well. The structural similarity [20] was defined based on the graph-edit-distance method [20,21]. Graph-edit-distance is the minimum required operations to alter the structure of one hierarchy to another.

Node deletion and node substitution are two metrics used to derive the graph-edit-distance. In comparison of two WBSs, mapping nodes with a higher degree of similarity requires less effort than mapping nodes with lower similarities. In addition, eliminating the nodes which are not mapped requires an extra effort. Thus, a smaller required effort to change the structures of two WBSs can result in a higher structural similarity. Node deletion measures the required effort to eliminate unmapped nodes. Node substitution is the required effort to link the mapped nodes.

Equations 12 and 13 determine the deletion and substitution efforts between $WBS^{L_1}_{N_1}$ and $WBS^{L_2}_{N_2}$. Equation 12 defines deletion effort (DE) as a ratio of the number of unmapped nodes ($|UN|$) over the total number of nodes. Substitution effort (SE) is calculated in Equation 13 by averaging dissimilarity of the mapped nodes. In other words, a higher degree of similarity between the nodes will result in a lower substitution effort. Arithmetic average of deletion and substitution efforts produces a representative of structural dissimilarity and its complement was determined as a measure of structural similarity as shown in Equation 14.

$$DE\left(WBS^{L_1}_{N_1}, WBS^{L_2}_{N_2}\right) = \frac{|UN|}{|N_1| + |N_2|} \tag{12}$$

$$SE\left(WBS_{N_1}^{L_1}, WBS_{N_2}^{L_2}\right) = \frac{2 * \sum_{(n_i,m_j)\in mapped\ nodes}\left(1 - sim_{average}(n_i, m_j)\right)}{|N_1| + |N_2| - |UN|} \quad (13)$$

$$Structural\ similarity\ \left(WBS_{N_1}^{L_1}, WBS_{N_2}^{L_2}\right) = 1 - \frac{DE\left(WBS_{N_1}^{L_1}, WBS_{N_2}^{L_2}\right) + SE\left(WBS_{N_1}^{L_1}, WBS_{N_2}^{L_2}\right)}{2} \quad (14)$$

3 Experimental results

Three experts in the project management domain were asked to develop four WBSs for four different construction projects: a bridge construction (steel girder with composite concrete slab), a steel-framed office building, a reinforced concrete-framed residential building, and a road-widening project. Table 1 presents the 12 WBS samples created by these subject matter experts. These samples were utilized to evaluate the performance of the proposed metrics in distinguishing the projects and retrieving the most similar samples.

As was mentioned before, an equal threshold was used in comparison of parents and siblings. To determine the impact of this threshold on the proposed metrics, the results were explored by changing the threshold in the range of 0.5 to 0.8 with 0.05 intervals.

Table 1. Created samples by the experts.

Experts	Developed samples	Represented by
expert 1	$bridge\ construction_1$	B_1
	$concrete\ structure\ building_1$	C_1
	$steel\ structure\ building_1$	S_1
	$road\ maintenance_1$	M_1
expert 2	$bridge\ construction_2$	B_2
	$concrete\ structure\ building_2$	C_2
	$steel\ structure\ building_2$	S_2
	$road\ maintenance_2$	M_2
expert 3	$bridge\ construction_3$	B_3
	$concrete\ structure\ building_3$	C_3
	$steel\ structure\ building_3$	S_3
	$road\ maintenance_3$	M_3

3.1 Retrieval precision and recall

Any retrieved samples resulted from a search process will only fall in one of the following categories: “retrieved and relevant”, “retrieved and not relevant”, “not retrieved and relevant” or “not retrieved and not relevant” [22]. Recall is the proportion of the retrieved and relevant samples from all the relevant samples [22]. Precision is the proportion of the retrieved and relevant items from all retrieved items [22], as presented in Equation 15.

The samples B_1 , C_1 , S_1 and M_1 were chosen to query the database and compute the precision. In each test, one of these samples was compared with all the stored samples in order to retrieve the relevant samples. For instance, the relevant samples to B_1 are B_2 and B_3. In each test, the retrieval process continued until it satisfy the recall score. The results were obtained based on two recall scores of 0.5 and 1. The recall score equals to 0.5 when only one of the two stored relevant samples are retrieved, and it equals to one when both relevant samples are retrieved. For instance, the node similarity scores between B_1 and stored samples are sorted in Table 2. Given that B_1, B_2 and B_3 are developed for the same project, the most relevant sample to B_1 are B_2 and B_3. In this query, assuming recall score is equivalent to 1, all the relevant samples to B_1 (B_2 and B_3) must be retrieved. As a result, the retrieving precision will be equal to 0.67.

$$Retrieving\ precision = \frac{|\{Relevant\ samples\} \cap \{Retrived\ samples\}|}{|\{Retrieved\ samples\}|} \quad (15)$$

Table 2. Comparing B1 with stored samples with a threshold of 0.65

Rank	Query sample	Stored sample	Node similarity score
1	b1	b2	0.64
2	b1	s2	0.56
3	b1	b3	0.48
4	b1	c1	0.47
5	b1	s1	0.40
6	b1	c2	0.39
7	b1	s3	0.35
8	b1	c3	0.34
9	b1	m2	0.13
10	b1	m3	0.13
11	b1	m1	0.05

Figure 4 illustrates the retrieval precision for node similarity and structural similarity scores. It can be concluded that the structural similarity with a threshold between 0.7 to 0.75 results in the highest precision scores.

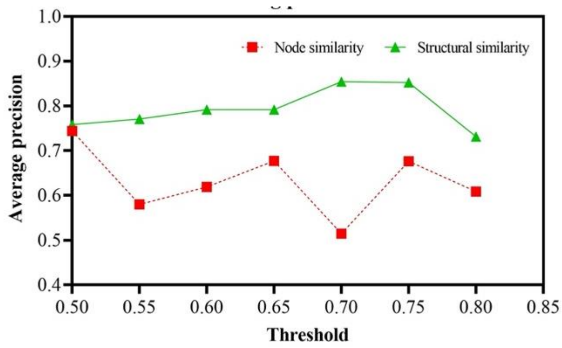


Figure 4. Average precision scores.

3.2 Properties of similarity measures

Symmetry and reflexivity are properties of a similarity measurement [23,24]. A similarity function S: S×S→ [0,1] on a set S, must fulfill two properties that are presented in Equations 16 and 17.

$$Sim(X,Y) = sim(Y,X)\ (Symmetry) \tag{16}$$

$$Sim(X,X) = 1\ (Reflexivity) \tag{17}$$

$$\forall\, X, Y \in S$$

3.2.1 Symmetry property of work breakdown structure similarity

Symmetry error, calculated using Equation 18, was defined to determine the symmetry fulfillment, in comparing WBS A and B. As illustrated in Figure 5, both node and structural similarity result in low degrees of error of symmetry property, namely the structural similarity error is close to zero.

$$Symmetry\ error = \frac{|sim(A,B) - sim(B,A)|}{average[sim(A,B), sim(B,A)]} \tag{18}$$

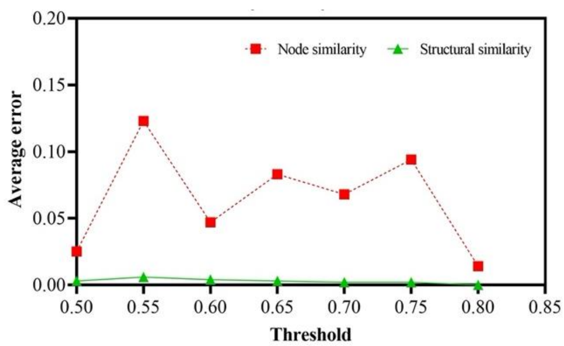


Figure 5. The average of the symmetry errors

3.2.2 Reflexivity property of work breakdown structure similarity

The reflexivity error, calculated using Equation 19, determines the compliance of the node and structural similarities with symmetry property by comparing a WBS to itself. It is evident from Figure 6 that both node and structural similarity result in very low levels of symmetry errors. Specifically, the error for the structural similarity is negligible, indicating better compliance of this metric with the symmetry property.

$$Reflexivity\ error = 1 - sim(A\,, A) \tag{19}$$

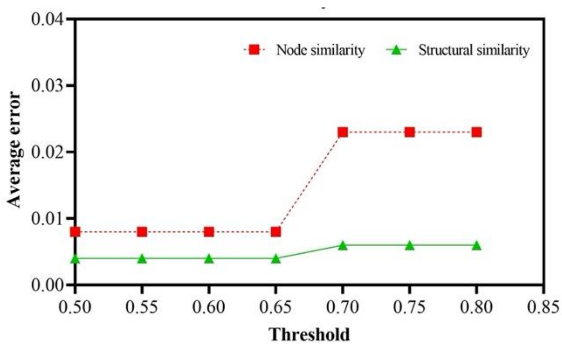


Figure 6. The average of the reflexivity errors

4 Conclusion

A quantitative similarity measurement of construction projects can improve the data retrieval process in knowledge management and case-based reasoning systems. This research utilized the WBS to measure the similarity amongst various construction projects. The WBS theoretically encompasses a significant portion of the tasks and services required to accomplish the project objectives. Thus, the WBS can be used as a comprehensive tool to distinguish constriction projects. The proposed method can be utilized to find the most similar projects and help the project managers use the relevant information and documents of the retrieved projects. This method can also be used as a main metric in exiting knowledge retrieval systems, such as case-based reasoning.

Node and structural similarity metrics were proposed to measure the similarity of construction projects using their WBSs. The results from applying these metrics indicate that the structural similarity performs better than the node similarity in retrieving the relavant samples. Additionally, the results show that the proposed metrics are in compliance with similarity measurement properties.

5 Limitations and future works

To ensure that the work breakdown structures encompass the entire scope of a project, the sample WBSs were developed based on the project management institute guidelines [11]. This approach might not be fully followed by the practitioners in all construction projects which can limit the performance of the proposed method.

The insufficient number of test samples was another major limitation of this research. The performance of the system can be investigated on a larger sample size. Another limitation of this research was the utilized lexical database, which is WordNet. This is a generic database for English words, in which some of the technical words in the construction domain are not accurately defined. For instance, the word "rebar", which in construction is recognized as reinforcement steel, is not defined in WordNet. This issue can hinder the performance of the metrics in comparing tasks. Thus, there is an opportunity to develop and utilize a semantic network, similar to WordNet, for the construction domain.

6 References

[1] Tserng, H. P., & Lin, Y. C., Developing an activity-based knowledge management system for contractors. Automation in construction, 13(6), 781-802, 2004.

[2] Alavi, M., & Leidner, D. E, Knowledge management and knowledge management systems: Conceptual foundations and research issues. MIS quarterly, 107-136, 2001.

[3] Hahn, J., & Subramani, M., A framework of knowledge management systems: issues and challenges for theory and practice. ICIS 2000 Proceedings, 28 2000.

[4] Tan, H. C., Carrillo, P. M., Anumba, C. J., Bouchlaghem, N., Kamara, J. M., & Udeaja, C. E., Development of a methodology for live capture and reuse of project knowledge in construction. Journal of management in engineering, 23(1), 18-26, 2007.

[5] Kolodner, J. L., An introduction to case-based reasoning. Artificial intelligence review, 6(1), 3-34, 1992.

[6] An, S. H., Kim, G. H., & Kang, K. I., A case-based reasoning cost estimating model using experience by analytic hierarchy process. Building and Environment, 42(7), 2573-2579, 2007.

[7] Raphael, B., Domer, B., Saitta, S., & Smith, I. F., Incremental development of CBR strategies for computing project cost probabilities. Advanced Engineering Informatics, 21(3), 311-321, 2007.

[8] Goh, Y. M., & Chua, D. K. H., Case-based reasoning approach to construction safety hazard identification: Adaptation and utilization. Journal of Construction Engineering and Management, 136(2), 170-178, 2010.

[9] Ryu, H. G., Lee, H. S., & Park, M., Construction planning method using case-based reasoning (CONPLA-CBR). Journal of Computing in Civil Engineering, 21(6), 410-422, 2007.

[10] Mikulakova, E., König, M., Tauscher, E., & Beucke, K., Knowledge-based schedule generation and evaluation. Advanced Engineering Informatics, 24(4), 389-403, 2010.

[11] PMBOK® Guide. Sixth Edition, Project Management Institute, 2017.

[12] Chowdhury, G.G., Natural language processing. Annual review of information science and technology, 37, 51-89, 2003.

[13] Gomaa, W. H., & Fahmy, A. A., A survey of text similarity approaches. International Journal of Computer Applications, 68(13), 13-18, 2013.

[14] Landauer, T. K., & Dumais, S. T., A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. Psychological review, 104(2), 211, 1997.

[15] Sowa, J. F., Semantic networks. John_Florian_Sowa isi [2012-04-20 16: 51]> Author [2012-04-20 16: 51], 2012.

[16] Miller, G. A., WordNet: a lexical database for English. Communications of the ACM, 38(11), 39-41, 1995.

[17] Meng, L., Huang, R., & Gu, J., A review of semantic similarity measures in wordnet. International Journal of Hybrid Information Technology, 6(1), 1-12, 2013.

[18] Wu, Z., & Palmer, M., Verb semantics and lexical selection. arXiv preprint cmp-lg/9406033, 1994.

[19] Mihalcea, R., Corley, C., & Strapparava, C., Corpus-based and knowledge-based measures of text semantic similarity. In Aaai (Vol. 6, No. 2006, pp. 775-780), 2016.

[20] Dijkman, R., Dumas, M., Van Dongen, B., Käärik, R., & Mendling, J., Similarity of business process models: Metrics and evaluation. Information Systems, 36(2), 498-516, 2011.

[21] Hart, P. E., Nilsson, N. J., & Raphael, B., A formal basis for the heuristic determination of minimum cost paths. IEEE transactions on Systems Science and Cybernetics, 4(2), 100-107, 1968.

[22] Buckland, M., & Gey, F., The relationship between recall and precision. Journal of the American society for information science, 45(1), 12-19, 1994.

[23] Ehrig, M., Koschmider, A., & Oberweis, A., Measuring similarity between semantic business process models. In Proceedings of the fourth Asia-Pacific conference on Comceptual modelling-Volume 67 (pp. 71-80), 2007.

[24] Richter, M. M., Classification and learning of similarity measures. In Information and

Classification (pp. 323-334). Springer, Berlin, Heidelberg, 1993.