

Research Article

Matching Transportation Ontologies with Word2Vec and Alignment Extraction Algorithm

Xingsi Xue ^{1,2}, Haolin Wang,^{1,2} Jie Zhang,³ Yikun Huang ⁴, Mengting Li,² and Hai Zhu ⁵

¹Intelligent Information Processing Research Center, Fujian University of Technology, Fuzhou, Fujian 350118, China

²School of Computer Science and Mathematics, Fujian University of Technology, Fuzhou, Fujian 350118, China

³School of Computer Science and Engineering, Yulin Normal University, Yulin, Guanxi 537000, China

⁴Concord University College Fujian Normal University, Fuzhou, Fujian 350117, China

⁵School of Network Engineering, Zhoukou Normal University, Zhoukou, Henan 466001, China

Correspondence should be addressed to Yikun Huang; fjnuhyk@163.com

Received 10 April 2021; Accepted 4 May 2021; Published 11 May 2021

Academic Editor: Chi-Hua Chen

Copyright © 2021 Xingsi Xue et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The development of intelligent transportation systems (ITSs) faces the challenge of integrating data from multiple unrelated sources. As one of the core technologies of knowledge integration in ITS, an ontology typically provides a normative definition of transportation domain that can be used as a reference for information integration. However, due to the subjectivity of domain experts, a concept may be expressed in multiple ways, yielding the ontology heterogeneity problem. Ontology matching (OM) is an effective method of addressing it, which is of help to further realize the mutual communication between the ontology-based ITSs. In this work, we first propose to use Word2Vec to model the entities in vector space and calculate their similarity values. Then, a stable marriage-based alignment extraction algorithm is presented to determine high-quality alignment. In the experiment, the performance of the proposal is tested by using the benchmark track of OAEI and real transportation ontologies. The experimental results show that our approach is able to obtain higher quality alignment results than OAEI's participants and other state-of-the-art ontology matching techniques.

1. Introduction

Data in the transportation domain are complex and varied [1–3]. These data come from a variety of data collection methods such as traffic sensors, surveys, and devices [4]. Therefore, the development of intelligent transportation systems (ITSs) faces the challenge of integrating data from multiple unrelated sources [5–7]. These data are semantically imprecise, conceptually ambiguous, and informative. As one of the core technologies of knowledge integration in ITS, an ontology typically provides a formal and normative definition of domain knowledge [8–10]. They enable collaboration between ontology-based ITSs by defining related concepts in the domain and the relationships between concepts [4]. However, due to the subjectivity of domain experts, the notion of a concept might be expressed in multiple ways [11, 12]. In order to achieve mutual

communication between ontology-based ITSs, it is important to determine the logical relationships between heterogeneous ontologies [13]. Ontology matching (OM) is an important technique to solve the problem of semantic heterogeneity [14, 15], which is dedicated to discovering correspondences between related entities (e.g., classes and properties) in different ontologies [16]. For this reason, it is effective to use ontology matching techniques to solve the existing semantic heterogeneity of transport ontologies.

In recent years, researchers have proposed a large number of ontology alignment strategies and developed various semiautomated or automated ontology matching systems [17–19]. However, the existing ontology matching schemes have many drawbacks: the matcher's poor ontology similarity calculation, inefficient extraction of ontology mapping results, etc. To address these drawbacks, we first propose to use Word2Vec [20] to model the entities in vector

space, calculate their similarity values, and use Wikipedia training data to improve the model's generalizability and the alignment's quality. Moreover, a stable marriage-based ontology extraction algorithm is presented to improve the quality of alignment.

The rest content of this article is as follows. In Section 2, we briefly describe the application of the ontology matching to transportation data. Section 3 presents ontology and ontology matching in detail. In Section 4, we develop an ontology matching system using the Word2Vec model. We propose Word2Vec-based similarity measure and stable marriage-based ontology extraction algorithm in Sections 5 and 6. In Section 7, experimental results and analysis are described. Finally, conclusion and future work are provided in Section 8.

2. Related Work

Ontology matching is well suited to solve the problems arising from semantic ambiguity and large data volume in transportation data [21–23]. Benvenuti et al. [24] integrated Transmodel ontologies and KPIOnto to facilitate the study of public road monitoring systems. The Transmodel is a reference data model about the European public transport information system that represents traffic ontologies and their relationships. Bermejo et al. [25], in order to avoid the use of a central decision point in a traffic network, proposed to treat each vehicle as an ontology that gives it reasoning capabilities. In emergency traffic control, each vehicle in the proposed system is a decision point that considers the state and location of neighboring vehicles and collaborates with them to reach consensus in real time. Overall, ontology provides a prescriptive approach to the development of knowledge in the transportation domain that can support integrated information in a variety of ways. Standardization efforts in transportation can be greatly assisted by the ontology engineering approach.

Recently, the technique for calculating words' similarity and relevance using the Word2Vec model from word embedding has become a research hotspot and is gradually applied to the domain of ontology matching. Xue and Pan [26] modeled the ontology in vector space and then used the linguistic information of the entities to reduce the dimensionality, which improved the efficiency of similarity calculation and entity matching. Zhang et al. [27] introduced word embedding techniques to the field of ontology matching and proposed a hybrid method that incorporates word embedding into the calculation of semantic similarity between elements. Teslya and Savosin [28] proposed a Word2Vec vector-based language model for the ontology mapping problem. The model extends on the basis of specific ontology relations. The semantics of the language is used to match ontologies without considering the form of words or specific terms. It can be seen that it is feasible to calculate the semantic similarity using the Word2Vec model. As for alignment result extraction, with a large number of matching methods being proposed, researchers usually need to integrate multiple strategies to improve the quality of alignment. To better extract matching results, a stable marriage-based

ontology extraction algorithm is proposed in this work, which further improves the performance of the matcher.

3. Ontology and Ontology Matching

An ontology is a conceptualized normative description on the domain knowledge [29–31]. Specifically, an ontology normatively defines classes, properties, other entities in a domain [32, 33], and the relationships between them. Figure 1 shows the ontology in a road accident [34]. The words in the rectangular box are classes, e.g., “Ting,” “Vehicle,” and “Insurance Company.” The hollow arrows indicate the structural relationship between two classes, e.g., “Official Agency” is a subclass of “Insurance Company” and “Road Accident” is a subclass of “Event.” The black arrows represent properties that describe the relationship between the two classes. However, the same entity could be constructed in multiple ways in different ontologies, yielding the problem of semantic heterogeneity between ontologies.

In order to illustrate the matching problem, the result of matching between two simple ontologies O and O' is presented in Figure 2. The two ontologies in the figure have descriptions of classes, properties, and instances. Classes are displayed in rectangles. Structure-based relationships are shown as broken line arrows. In O , “Chairman” is a specialization (subclass) of “Person.” Correspondences are shown as blue double arrows that connect classes from O to classes from O' and depict their relationships. There are symbols: \perp , \subseteq (or \supseteq), and \equiv , which mean disjointness, more specific (or less specific), and equivalence relation, respectively. For example, the “Subject Area” in one ontology is equivalent to the “Topic” in another ontology, and the “Regular Author” is an irrelevant relationship with the “Reviewer.”

4. The Framework of W2V-OM

This work constructs an ontology matcher (W2V-OM) that calculates the similarity values of two entities using the Word2Vec model, as shown in Figure 3. The dataset for training the Word2Vec model is the English Language Wikipedia articles in the Wikipedia database [35]. The corpus is universal and can cope with language processing problems in many domains. These textual data are unstructured and need to be preprocessed into structured data. After the model is trained, the source transportation ontology and the target ontology are parsed. The entities extracted from the ontology are fed into the word2vec model to calculate the cosine similarity and integrate the linguistic-based similarity measure to yield the similarity matrix. Then, the ontology mapping results are obtained using the stable marriage-based ontology extraction algorithm. Finally, the ontology matching quality is evaluated based on reference alignment.

5. Word2Vec-Based Similarity Measure

The similarity measure is a function where the information of two ontology entities is used as input and a real number between [0, 1] is output to represent their similarity [36]. Specifically, the closer the result is to 1, the more similar they

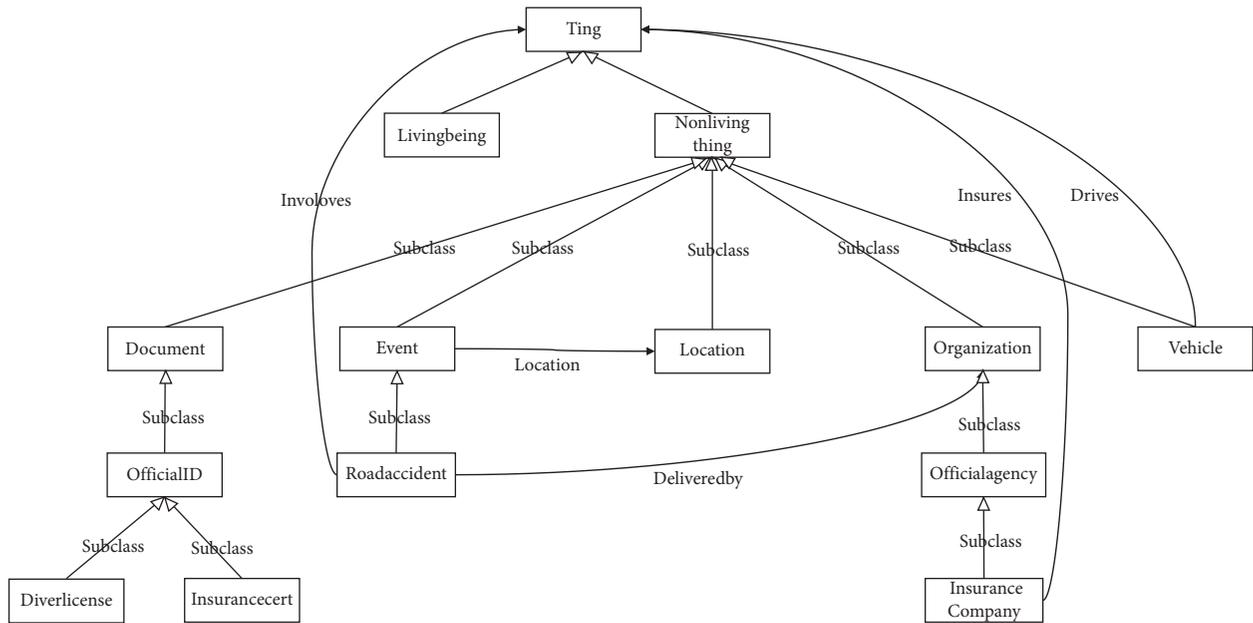


FIGURE 1: An example of a road accident ontology.

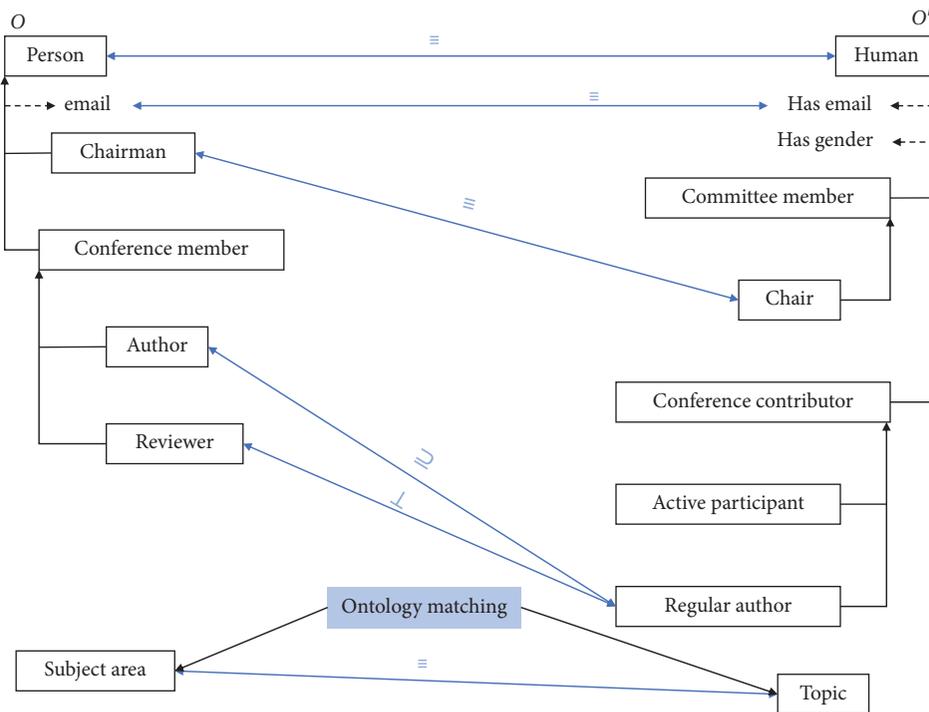


FIGURE 2: An example of ontology alignment.

are; the closer the result is to 0, the less similar they are. Similarity measure is an important part of the ontology matching process. Utilizing different similarity measure affects the results of ontology alignment. In this work, we use two categories of similarity measures to calculate the similarity values of two entities, i.e., linguistic-based measure and cosine similarity measure using the Word2Vec model. Word2Vec is a language model Natural Language Processing (NLP) where words or phrases are represented as real

number vectors. Similar words usually have the proximity of vectors and are mapped to the same region, as shown in Figure 4. With regard to the ontology representation in vector space, it means that a class or a property of ontologies can be represented in dimensions of the vector space. Specifically, the different classes or properties are uniquely represented in the vector space. The vector space covers all classes and properties in both ontologies. In this work, the dimensions of the vector space are determined by all the classes and properties in the

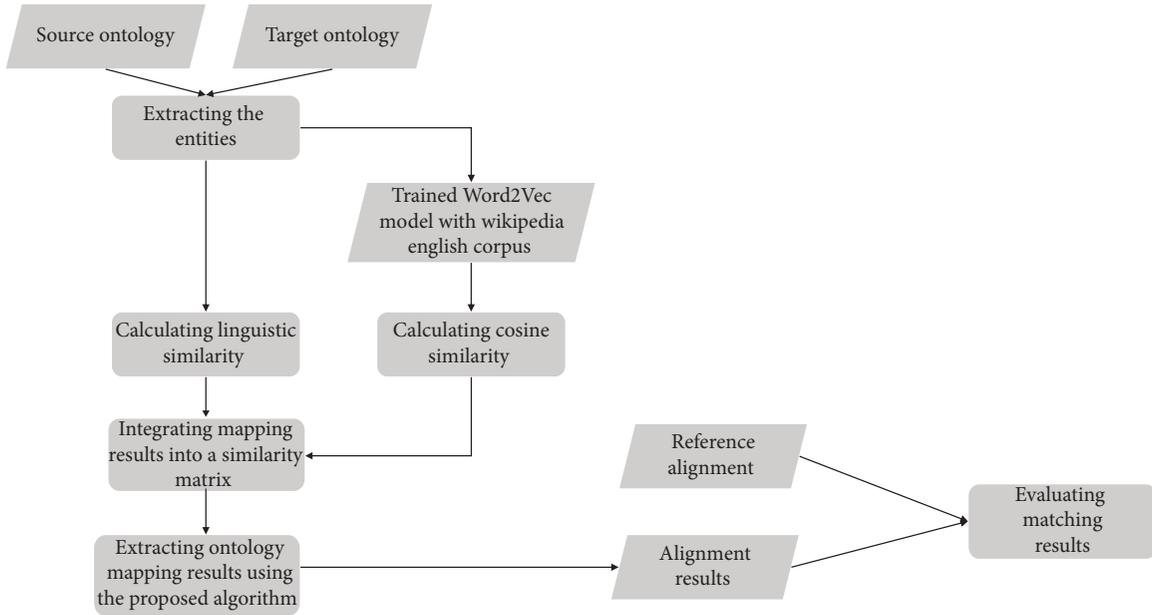


FIGURE 3: The framework of W2V-OM.

two ontologies. The Word2Vec model is trained using the Wikipedia English corpus. Each entity is represented as a vector in vector space, and then, the similarity of the two entities is calculated using the cosine similarity formula. The formula is defined as follows:

$$\text{Cosine Similarity}(V_{w_1}, V_{w_2}) = \frac{V_{w_1} \cdot V_{w_2}}{V_{w_1} \cdot V_{w_2}}, \quad (1)$$

where V_{w_1} and V_{w_2} are, respectively, the vectors of two words w_1 and w_2 and V_{w_1} and V_{w_2} , respectively, denote their norms.

The linguistic similarity between two words is calculated by semantic relations (synonymy and antonymy), which is generally done using dictionaries and lists of synonyms. WordNet [37], a vocabulary database that builds semantic networks based on the semantic information of words, is used to calculate similarity. The linguistic similarity of two words w_1 and w_2 is 1 when w_1 and w_2 are synonyms in WordNet; the similarity is 0.5 when w_1 and w_2 are hypernym in WordNet; in other cases, the similarity is 0.

The two similarity measures produce two similarity matrices, and it is necessary to use an aggregation strategy to set the different matrices into one matrix. In this work, we empirically use the maximum strategy to integrate the similarity measures, i.e., the larger one of two similarity values is selected as the final similarity value, which is of help to ensure the completeness of the alignment.

6. Stable Marriage-Based Alignment Extraction

Integrating the results computed from the similarity measures in a similarity matrix. The i th row and j th column of this matrix represent the entities e_{S_i} and e_{T_j} in the source ontology O_S and the target ontology O_T , respectively. The values in the matrix indicate the similarity of the two entities. Larger similarity values indicate higher confidence in the

equivalence of the two entities and vice versa indicates less confidence. In this paper, we propose a stable marriage-based ontology extraction algorithm that incorporates a thresholding strategy to obtain better mapping results. The specific steps are as follows: (1) all similarity values in the matrix are sorted in descending order, (2) record the position $(e_{S_i}, e_{T_j}, m_{ij})$ of the maximum similarity in the matrix, where m_{ij} is the maximum similarity, (3) set the value in the same row and column of m_{ij} as 0, and (4) repeat the above three steps until all similarity values in the matrix are 0.

Figure 5 presents the results of extracting the ontology mapping using the proposed method. As shown in the figure, six entity correspondences were finally extracted, which are $(e_{S,1}, e_{T,1}, 0.95)$, $(e_{S,2}, e_{T,2}, 0.88)$, $(e_{S,3}, e_{T,3}, 0.6)$, $(e_{S,5}, e_{T,5}, 0.6)$, $(e_{S,6}, e_{T,5}, 0.6)$, and $(e_{S,4}, e_{T,4}, 0.1)$. The proposed algorithm terminates when all similarity values in the matrix are zero, which may result in extracting some entity correspondences with low similarity. For the mapping results, these low similarities are noise. This work therefore incorporates a threshold strategy. A threshold parameter is set and the algorithm is terminated when all values in the similarity matrix are less than the threshold. Assume a threshold of 0.5, i.e., a similarity of less than 0.5 is not reliable. Then, the similarity matrix extraction results are $(e_{S,1}, e_{T,1}, 0.95)$, $(e_{S,2}, e_{T,2}, 0.88)$, $(e_{S,3}, e_{T,3}, 0.6)$, $(e_{S,5}, e_{T,5}, 0.6)$, and $(e_{S,6}, e_{T,5}, 0.6)$.

7. Experiment

7.1. Experimental Configuration. In the experiment, the performance of our proposal is tested using real sensor ontologies in the transportation field, as well as the benchmark track provided by the Ontology Alignment Evaluation Initiative (OAEI). The benchmark test library is constructed from reference ontologies in different domains. Each test case in the benchmark track contains two

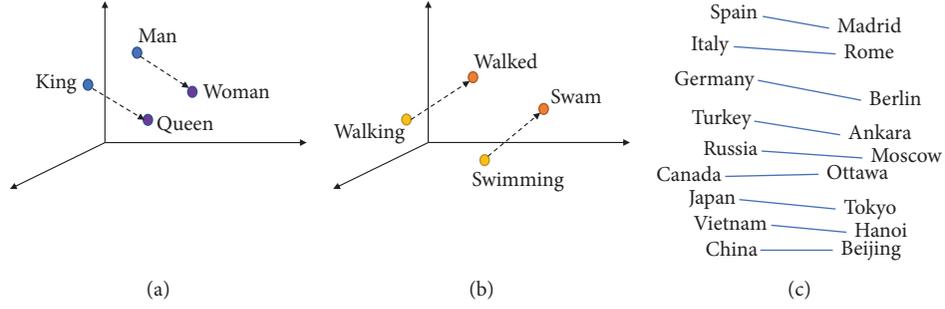


FIGURE 4: Model the words in vector space. (a) Male-Female. (b) Verb tense. (c) Country-Capital.

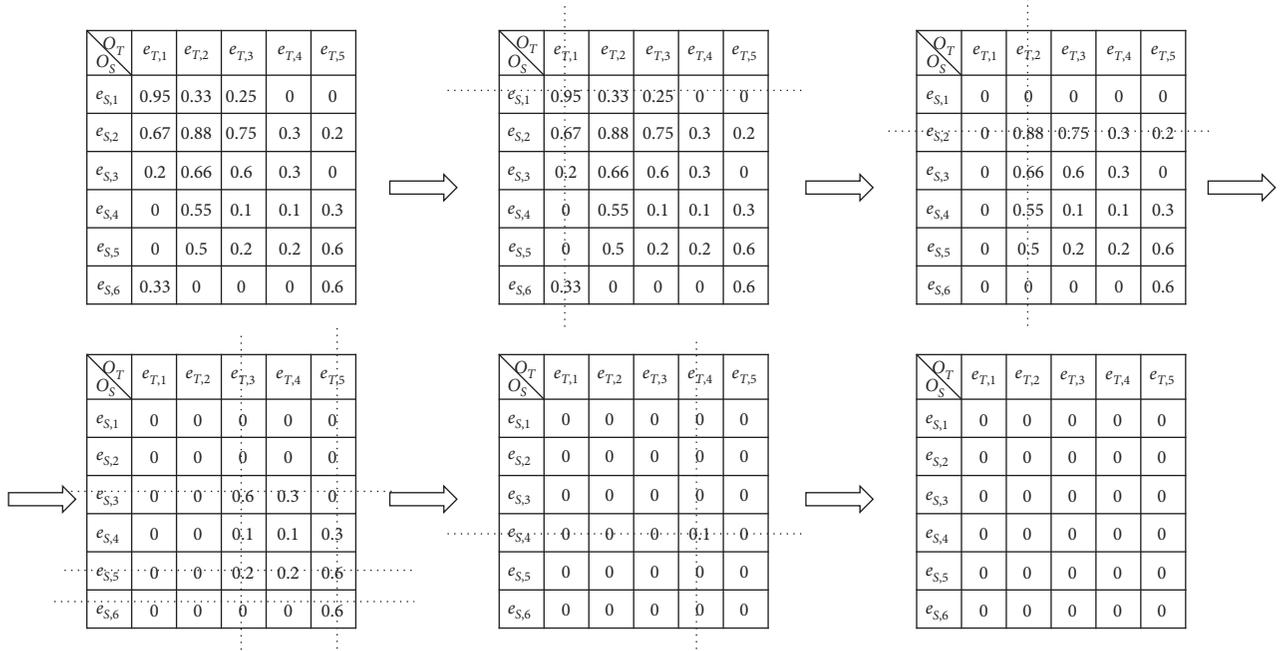


FIGURE 5: An example of alignment extraction.

ontologies to be matched (a target ontology and a source ontology) and a reference alignment for evaluating the effectiveness of the ontology matcher. The real sensor ontologies used are OSSN, SN, SOSA, and SSN. Table 1 presents a detailed description of the benchmark test cases, and the concise introduction of the sensor ontologies is given in Table 2. In order to assess the quality of ontology matching results, the following are the traditional definitions of ontology alignment metrics:

$$recall = \frac{\text{correct_found_correspondences}}{\text{all_possible_correspondences}},$$

$$precision = \frac{\text{correct_found_correspondences}}{\text{all_found_correspondences}}, \quad (2)$$

$$f\text{-measure} = \frac{2 * recall * precision}{recall + precision},$$

where *precision* and *recall* denote the accuracy and completeness of the alignment results, respectively, and *f-measure* is the harmonic mean of *precision* and *recall* to balance them.

7.2. Comparison with OAEI's Participants. Figures 6–8 present the comparison between W2V-OM and the participants of OAEI in terms of recall, precision, and f-measure, respectively. In the figures, the horizontal axis indicates the testing case ID, the vertical axis indicates the alignment's evaluation metric, and the legends indicate the different matching systems. As shown in the figure, the W2V-OM is higher than the other OAEI's participants in terms of recall and F-measure. With respect to precision, our approach outperforms the other participants in most cases. In summary, the performance of W2V-OM proposed in this work is better than OAEI's participants and can determine high-quality ontology alignment.

TABLE 1: A detailed description of the benchmark test case.

Cases ID	Case introduction
101-104	Two same ontologies
201-210	Two ontologies with different linguistic and lexical features
221-247	Two different structures of ontologies
248-262	Two ontologies with different structure, linguistic, and lexical features

TABLE 2: A concise introduction of the sensor ontologies.

Ontology name	Scale	Website
Semantic Sensor Network (SSN) ontology	55 entities	https://www.w3.org/ns/ssn/
Original Semantic Sensor Network (OSSN) ontology	107 entities	https://www.w3.org/2005/Incubator/ssn/wiki/SSN#Sensor
SensorOntology2009 (SN) ontology	152 entities	https://www.w3.org/2005/Incubator/ssn/wiki/SensorOntology2009
Sensor, Observation, Sample, and Actuator (SOSA) ontology	42 entities	https://www.w3.org/ns/sosa/

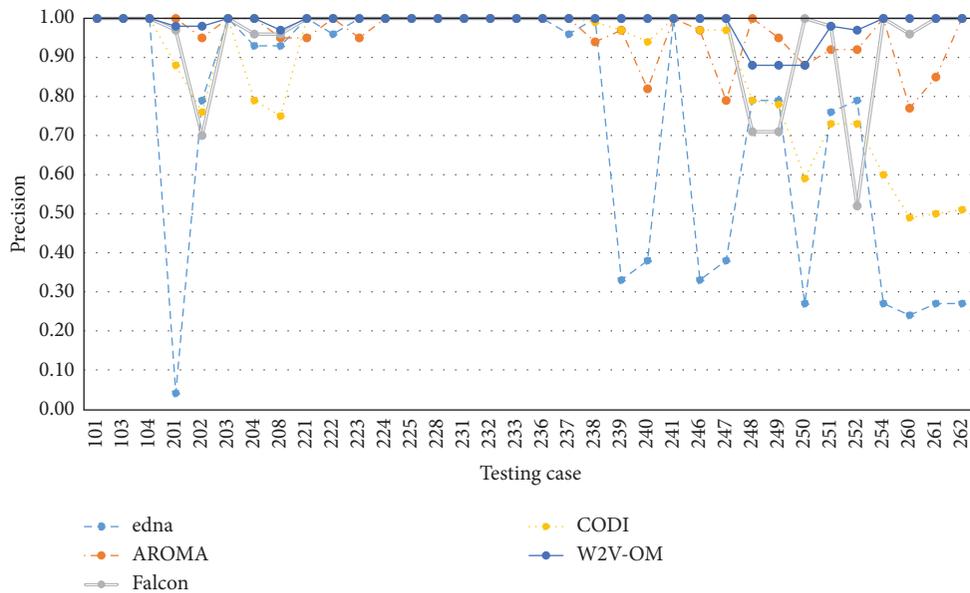


FIGURE 6: Comparison among W2V-OM and OAEI’s participants in terms of precision.

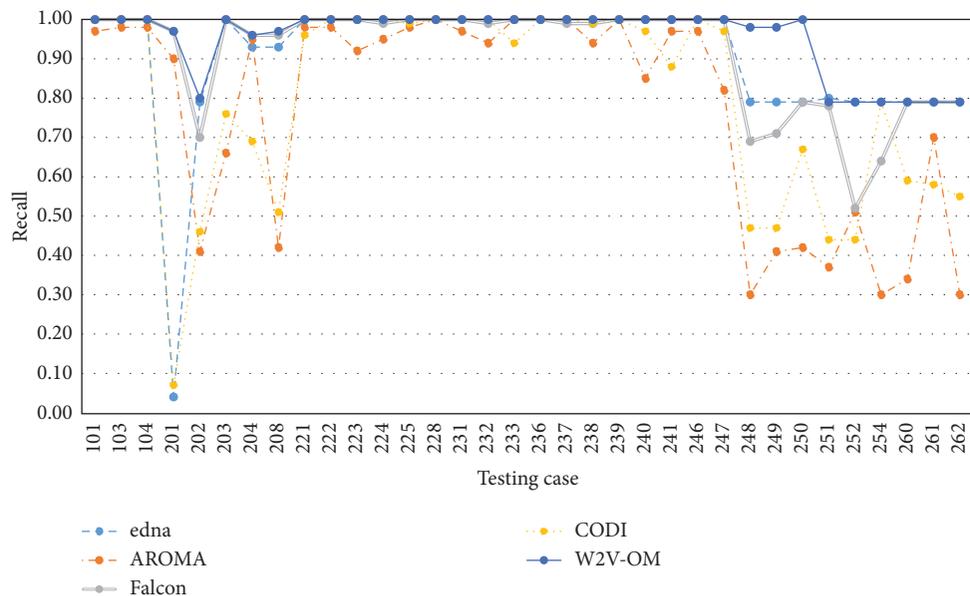


FIGURE 7: Comparison among W2V-OM and OAEI’s participants in terms of recall.

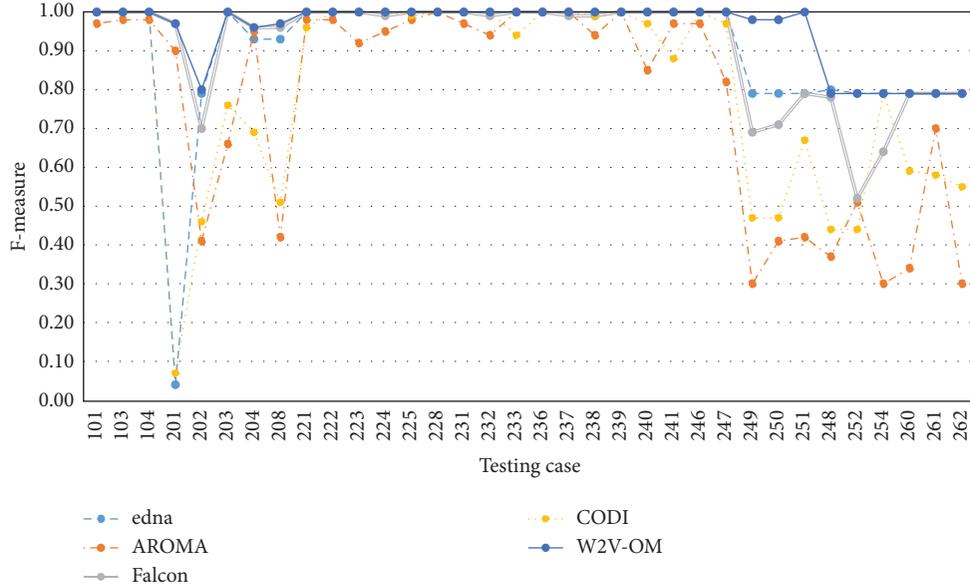


FIGURE 8: Comparison among W2V-OM and OAEI's participants in terms of F-measure.

TABLE 3: Comparison among W2V-OM and the state-of-the-art ontology matchers.

Alignment task	Ontology alignment metrics	WordNet-based matcher	SF-based matcher	Jaro-Winkler-based matcher	Levenshtein-based matcher	W2V-OM
SOSA-SN	Recall	1.00	0.07	1.00	1.00	1.00
	Precision	0.33	0.13	0.75	0.75	1.00
	f-measure	0.50	1.00	0.86	0.86	1.00
SOSA-OSSN	Recall	1.00	0.50	1.00	1.00	1.00
	Precision	0.67	0.20	1.00	1.00	1.00
	f-measure	0.80	0.29	1.00	1.00	1.00
SSN-OSSN	Recall	0.97	0.35	1.00	1.00	1.00
	Precision	0.80	0.06	0.94	1.00	1.00
	f-measure	0.88	0.11	0.97	1.00	1.00
SSN-SN	Recall	1.00	0.56	1.00	1.00	1.00
	Precision	0.52	0.02	0.90	1.00	1.00
	f-measure	0.70	0.04	0.95	1.00	1.00

7.3. Comparison with State-of-the-Art Ontology Matchers. Regarding the alignment of sensor ontologies, four popular ontology matchers were used as comparison groups, which are based on WordNet similarity [37], similarity flooding (SF) [38], Jaro-Winkler distance [39], and Levenshtein distance [40]. Table 3 shows the experimental results of sensor ontologies alignment. As can be seen from experimental results, W2V-OM outperforms other methods in four real sensor ontology matching tasks, which demonstrates the effectiveness of our approach.

Since our approach uses Word2Vec to map ontologies into vector space, the similarity is derived by calculating the vector cosine angle of the two entities. The model fully considers the string-based similarity measure and obtains a high similarity. In the mapping result extraction process, the similarity value of the matched entities should be the largest in the same row and column, which means that these two entities are the best alignment. The performance of the matcher is further improved by retaining the entity

correspondences with larger similarity values from the similarity matrix using the stable marriage strategy. To sum up, comparison with other matchers demonstrates the effectiveness of the proposed method.

8. Conclusion and Future Work

The purpose of matching transportation ontologies is to determine all the heterogeneous entity pairs. To this end, this work first models entities in vector space with Word2vec and uses the cosine similarity measure to calculate two entities' similarity value. After that, a stable marriage-based alignment extraction algorithm is used to determine high-quality alignment. The experimental results indicate that our approach can obtain higher quality alignment results compared to state-of-the-art ontology matchers and OAEI's participants.

In the future, we will adopt more advanced similarity measures to improve ontology similarity results. We also want to extend ontologies in the transportation domain,

such as the road traffic management ontology and the road accident ontology. Since transportation ontology matching requires particular alignment and knowledge background, specific techniques and strategies need to be proposed to enhance the quality of matching.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the Natural Science Foundation of Fujian Province (no. 2020J01875), National Natural Science Foundation of China (nos. 61801527 and 61103143), Fujian Province 13th Five-Year Plan Teaching Reform Project in 2019 (no. FBJG20190156), The Third Batch of Key Lifelong Education Projects in Fujian Province (no. ZS20033), 2018 Program for Outstanding Young Scientific Researcher in Fujian Province University the Research Innovation Team of Concord University College Fujian Normal University in 2020 (no. 2020-TD-001), and Scientific Research Project of Concord University College of Fujian Normal University in 2020 (no. KY20200203).

References

- [1] Y. Liu, X. Weng, J. Wan, X. Yue, H. Song, and A. V. Vasilakos, "Exploring data validity in transportation systems for smart cities," *IEEE Communications Magazine*, vol. 55, no. 5, pp. 26–33, 2017.
- [2] F. Ali, D. Kwak, P. Khan et al., "Transportation sentiment analysis using word embedding and ontology-based topic modeling," *Knowledge-Based Systems*, vol. 174, pp. 27–42, 2019.
- [3] L. Zhu, F. R. Yu, Y. Wang et al., "Big data analytics in intelligent transportation systems: a survey," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 1, pp. 383–398, 2018.
- [4] M. Katsumi and M. Fox, "Ontologies for transportation research: a survey," *Transportation Research Part C: Emerging Technologies*, vol. 89, pp. 53–82, 2018.
- [5] C.-H. Chen, F.-J. Hwang, and H.-Y. Kung, "Travel time prediction system based on data clustering for waste collection vehicles," *IEICE Transactions on Information and Systems*, vol. E102.D, no. 7, pp. 1374–1383, 2019.
- [6] C.-H. Chen, "An arrival time prediction method for bus system," *IEEE Internet of Things Journal*, vol. 5, no. 5, pp. 4231–4232, 2018.
- [7] C. H. Chen, F. Song, F. J. Hwang, and L. Wu, "A probability density function generator based on neural networks," *Physica A: Statistical Mechanics and Its Applications*, vol. 541, pp. 1–10, 2020.
- [8] C.-H. Chen, "A cell probe-based method for vehicle speed estimation," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E103.A, no. 1, pp. 265–267, 2020.
- [9] C. Jiang and X. Xue, "A uniform compact genetic algorithm for matching bibliographic ontologies," *Applied Intelligence*, pp. 1–16, 2021.
- [10] X. Xue and Y. Wang, "Ontology alignment based on instance using NSGA-II," *Journal of Information Science*, vol. 41, no. 1, pp. 58–70, 2015.
- [11] J. S. Pan, P. C. Song, S. C. Chu, and Y. J. Peng, "Improved compact cuckoo search algorithm applied to location of drone logistics hub," *Mathematics*, vol. 8, no. 3, pp. 1–19, 2020.
- [12] X. Xue and J. Zhang, "Matching large-scale biomedical ontologies with central concept based partitioning algorithm and adaptive compact evolutionary algorithm," *Applied Soft Computing*, vol. 106, pp. 1–11, 2021.
- [13] H. Liu, Y. Wang, and N. Fan, "A hybrid deep grouping algorithm for large scale global optimization," *IEEE Transactions on Evolutionary Computation*, vol. 24, no. 6, pp. 1112–1124, 2020.
- [14] X. Xue, X. Wu, C. Jiang, G. Mao, and H. Zhu, "Integrating sensor ontologies with global and local alignment extractions," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 6625184, 10 pages, 2021.
- [15] X. Xue, "A compact firefly algorithm for matching biomedical ontologies," *Knowledge and Information Systems*, vol. 62, pp. 1–17, 2020.
- [16] X. Xue, C. Yang, C. Jiang et al., "Optimizing ontology alignment through linkage learning on entity correspondences," *Complexity*, vol. 2021, Article ID 5574732, 12 pages, 2021.
- [17] X. Xue and Y. Wang, "Using memetic algorithm for instance coreference resolution," *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 2, pp. 580–591, 2015.
- [18] X. Xue and Y. Wang, "Optimizing ontology alignments through a memetic algorithm using both matchfmeasure and unanimous improvement ratio," *Artificial Intelligence*, vol. 223, pp. 65–81, 2015.
- [19] X. Xue, H. Yang, J. Zhang, and D. Chen, "An automatic biomedical ontology meta-matching technique," *Journal of Network Intelligence*, vol. 4, no. 3, pp. 109–113, 2019.
- [20] K. W. Church, "Word2Vec," *Natural Language Engineering*, vol. 23, no. 1, pp. 155–162, 2017.
- [21] X. Xue, J. Lu, and J. Chen, "Using NSGA-III for optimising biomedical ontology alignment," *CAAI Transactions on Intelligence Technology*, vol. 4, no. 3, pp. 135–141, 2019.
- [22] S. Fan, Z. Hua, V. C. Storey, and J. L. Zhao, "A process ontology based approach to easing semantic ambiguity in business process modeling," *Data & Knowledge Engineering*, vol. 102, pp. 57–77, 2016.
- [23] J. C.-W. Lin, Y. Shao, Y. Djenouri, and U. Yun, "ASRNN: a recurrent neural network with an attention model for sequence labeling," *Knowledge-Based Systems*, vol. 212, p. 106548, 2021.
- [24] F. Benvenuti, C. Diamantini, D. Potena, and E. Storti, "An ontology-based framework to support performance monitoring in public transport systems," *Transportation Research Part C: Emerging Technologies*, vol. 81, pp. 188–208, 2017.
- [25] A. J. Bermejo, J. Villadangos, J. J. Astrain, and A. Córdoba, "Ontology based road traffic management," *Intelligent Distributed Computing VI*, vol. 446, pp. 103–108, 2013.
- [26] X. Xue and J. S. Pan, "An overview on evolutionary algorithm based ontology matching," *Journal of Information Hiding and Multimedia Signal Processing*, vol. 9, no. 1, pp. 75–88, 2018.
- [27] Y. Zhang, X. Wang, S. Lai et al., "Ontology matching with word embeddings," *Lecture Notes in Computer Science*, Springer, Berlin, Germany, pp. 34–45, 2014.

- [28] N. Teslya and S. Savosin, "Matching ontologies with Word2Vec-based neural network," *Computational Science and Its Applications-ICCSA 2019*, Springer, Berlin, Germany, pp. 745–756, 2019.
- [29] T. R. Gruber, "Toward principles for the design of ontologies used for knowledge sharing," *International Journal of Human-Computer Studies*, vol. 43, no. 5-6, pp. 907–928, 1995.
- [30] G. Bella, F. Giunchiglia, and F. McNeill, "Language and domain aware lightweight ontology matching," *Journal of Web Semantics*, vol. 43, pp. 1–17, 2017.
- [31] P. Ochieng and S. Kyanda, "Large-scale ontology matching," *ACM Computing Surveys*, vol. 51, no. 4, pp. 1–35, 2018.
- [32] L. Otero-Cerdeira, F. J. Rodriguez-Martinez, and A. Gómez-Rodríguez, "Ontology matching: a literature review," *Expert Systems with Applications*, vol. 42, no. 2, pp. 949–971, 2015.
- [33] P. Shvaiko and J. Euzenat, "Ontology matching: state of the art and future challenges," *IEEE Transactions on Knowledge and Data Engineering*, vol. 25, no. 1, pp. 158–176, 2011.
- [34] J. Barrachina, P. Garrido, M. Fogue et al., "VEACON: a Vehicular Accident Ontology designed to improve safety on the roads," *Journal of Network and Computer Applications*, vol. 35, no. 6, pp. 1891–1900, 2012.
- [35] L. Denoyer and P. Gallinari, "The wikipedia XML corpus," *International Workshop of the Initiative for the Evaluation of XML Retrieval*, pp. 12–19, Springer, Berlin, Germany, 2006.
- [36] W. Gao, M. R. Farahani, A. Aslam, and S. Hosamani, "Distance learning techniques for ontology similarity measuring and ontology mapping," *Cluster Computing*, vol. 20, no. 2, pp. 959–968, 2017.
- [37] C. Fellbaum, "WordNet," *Theory and Applications of Ontology: Computer Applications*, Springer, Berlin, Germany, pp. 231–243, 2010.
- [38] S. Melnik, H. Garcia-Molina, and E. Rahm, "Similarity flooding: a versatile graph matching algorithm and its application to schema matching," in *Proceedings of the 18th International Conference on Data Engineering*, pp. 117–128, San Jose, CA, USA, February 2002.
- [39] W. W. Cohen, P. Ravikumar, and S. E. Fienberg, "A comparison of string distance metrics for name-matching tasks," *IIWeb*, vol. 3, pp. 73–78, 2013.
- [40] V. I. Levenshtein, "Binary codes capable of correcting deletions, insertions, and reversals," *Soviet Physics Doklady*, vol. 10, no. 8, pp. 707–710, 1966.