Query-oriented Entity Spatial-temporal Summarization in Fuzzy Knowledge Graph*

Erhe Yang
Shaanxi Normal University
erhyang@snnu.edu.cn

Fei Hao†
Shaanxi Normal University & University of Exeter
feehao@gmail.com

Aziz Nasridinov
Chungbuk National University
aziz@chungbuk.ac.kr

Geyong Min
University of Exeter
G.Min@exeter.ac.uk

Doo-Soon Park
Soonchunhyang University
parkds@sch.ac.kr

ABSTRACT

Knowledge Graph (KG) is a relatively new concept that has garnered a lot of attention. Furthermore, the information in KG is frequently ambiguous and imprecise, necessitating the creation of a Fuzzy Knowledge Graph (FKG). FKG describes the imprecise information of the entity by employing the fuzzy value of predicates or objects. Entity summarization can extract the most concise and important information from lengthy descriptions of an entity. Existing work, however, focuses solely on entity summarization in KG while ignoring the fuzziness of entity relationships in FKG. Thus, this paper proposed an FFCA-based approach for query-oriented entity spatial-temporal summarization. Fuzzy Formal Concept Analysis (FFCA) is used to turn the FKG into the regular KG initially. The summarized RDF triples can then be obtained by combining the time-centric and location-centric triadic concepts from diverse FKGs. Finally, various template-based queries are designed for evaluating the performance of the proposed approach.

KEYWORDS

Fuzzy Knowledge Graph, Spatial-temporal Summarization, Fuzzy Formal Concept Analysis, Triadic Formal Concept Analysis

ACM Reference Format:
Erhe Yang, Fei Hao, Aziz Nasridinov, Geyong Min, and Doo-Soon Park. 2022. Query-oriented Entity Spatial-temporal Summarization in Fuzzy Knowledge Graph. In *The 37th ACM/SIGAPP Symposium on Applied Computing (SAC '22), April 25–29, 2022, Virtual Event,* . ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3477314.3506987

*Produces the permission block, and copyright information
†Corresponding author: Fei Hao; Email:feehao@gmail.com.

SAC '22, April 25–29, 2022, Virtual Event,

ACM ISBN 978-1-4503-8713-2/22/04...$15.00
https://doi.org/10.1145/3477314.3506987

1 INTRODUCTION

Knowledge Graph (KG) represents the resource on the web using the RDF (Resource Description Framework) triples, which is gaining popularity. KG has been used successfully in a variety of fields, such as entity alignment [5], entity summarization [4] and question answering [10], etc. However, the information in the real world is often uncertain and ambiguous. Fuzzy Knowledge Graph (FKG) has been constructed for representing the imprecise and uncertain information on the Web and research efforts have been made to build fuzzy knowledge graph. For instance, a fuzzy knowledge base was built from several documents [15]. A framework [13] by feature selection was presented for improving the construction efficiency of fuzzy knowledge base. Nevertheless, the huge entity descriptions frequently cause information overload for users.

Formal Concept Analysis (FCA) [3] is a valuable mathematical tool for describing the binary relations between objects and attributes. As its extension, Triadic Formal Concept Analysis (TFCA) [8] aims at clustering and modeling the features of the triadic data. FCA has been adopted to tackle the problem of entity summarization. For instance, Kim et al. [7] proposed an FCA-based approach KAFCA that can obtain the ranked RDF triples by the weights of extents of concepts in concept lattice. Tasnim et al. [11] proposed an FCA-based approach to capture the entity temporal evolution from different versions of a KG. Yang et al. [14] proposed an incremental entity summarization approach leveraging incremental FCA algorithm, which can address entity summarization in dynamic knowledge graph. Due to its dynamic nature, KG constantly changes as time elapses. Besides, the spatial changes of the entity can also be regarded as a dynamic form. Because the spatial information of the entity in different locations is changing, the evolution information of the entity can be discovered. However, to the best of our knowledge, there has been no related work on the dynamic entity summarization in FKG. As extension of the prior works, this paper proposes a TFCA-based approach of entity spatial-temporal evolution summarization in the Fuzzy Knowledge Graph. The main contributions of this paper are summarized as follows:

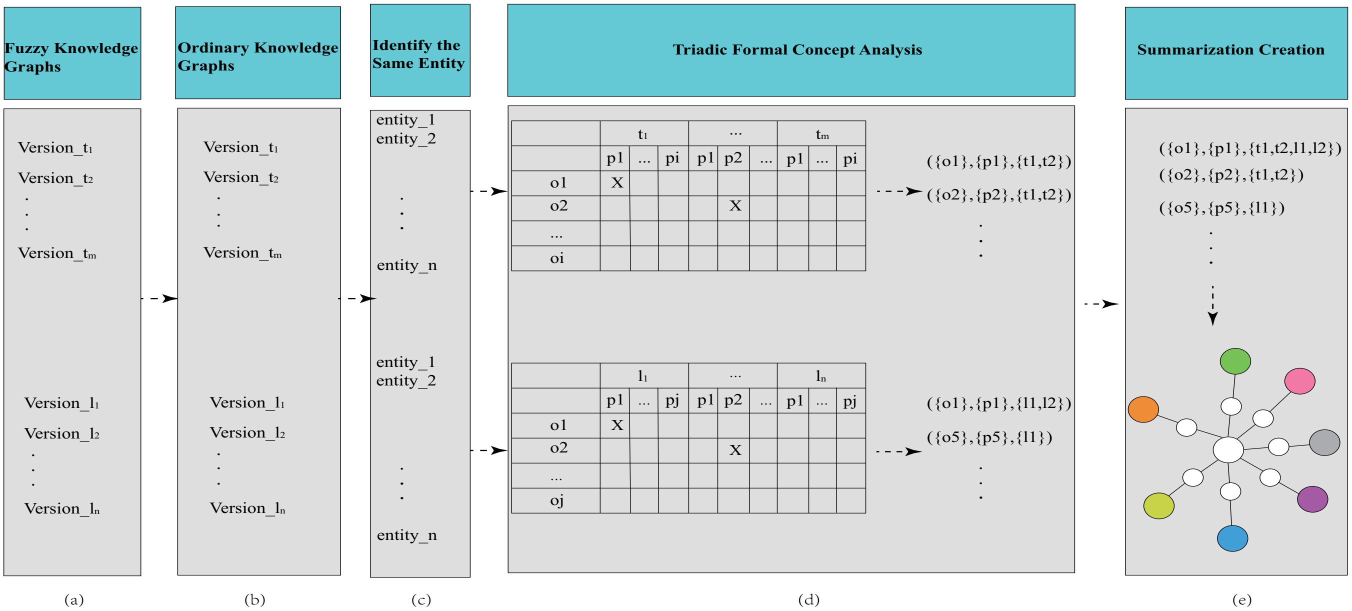


Figure 1: The framework of entity spatial-temporal evolution summarization in FKG.

- **Formalization of Entity Spatial-temporal Summarization in FKG**. We formulate the novel challenge of entity spatial-temporal evolution summarization in FKG to better mine the entity spatial-temporal evolution information.
- **TFCA-based Entity Spatial-temporal Evolution Summarization Approach**. For the given FKGs, we firstly transform the FKG into the ordinary KGs adopting Fuzzy Formal Concept Analysis. Then we obtain the entity spatial-temporal summarization adopting T-FCA. Finally, we integrate the obtained triadic concepts for generating the final summarization results.
- **Evaluation with Template-based SPARQL Queries**. For evaluating summarization results, we design several query templates for the proposed approach using the case study. We use the artificial data to assess the performance of the approach.

2 PRELIMINARIES AND PROBLEM DESCRIPTION

Definition 2.1. (**Temporal Entity Summarization**) Given the 4-tuple $< S_t, P_t, O_t, V >$ of the entity e from diverse versions of the FKG at time $T_1, T_2, T_3, ...$, temporal entity summarization is to generate a triple set (P_t, O_t, T), depicting that the entity e has the property (i.e., predicate and object) set at certain times.

Definition 2.2. (**Spatial Entity Summarization**) Given the 4-tuple $< S_l, P_l, O_l, V >$ of the entity e from diverse versions of the FKG at location $L_1, L_2, L_3, ...$, spatial entity summarization is to generate a triple set (P_l, O_l, L), depicting that the entity e has the property (i.e., predicate and object) set at certain locations.

We formulate the problem as follows.
(**Entity Spatial-temporal Summarization in Fuzzy Knowledge Graph**) Given that two 4-tuples $<S_t, P_t, O_t, V>$ and $<S_l, P_l, O_l, V>$ of the entity e from two groups of diverse versions of the FKG at time $T_1, T_2, T_3, ...$ and location $L_1, L_2, L_3, ...$, entity spatial-temporal evolution summarization is to generate a 4-tuple set (P, O, T, L), depicting that the entity e has the properties (i.e., predicate and object) set at certain times and locations.

3 PROPOSED APPROACH

3.1 Framework of Entity Spatial-temporal Evolution Summarization in FKG

To better understand the addressed problem, Fig. 1 shows the framework of the proposed approach. Firstly, as shown in Fig. 1 (a), various versions of FKG with time or location changes can be processed using FFCA for obtaining the ordinary knowledge graph shown in Fig. 1 (b). Our general idea is that we consider the vague value in FKG as "1" if the value greater than the given threshold, otherwise the value is deemed as "0". Then the two groups of the obtained molecules can be selected as shown in Fig. 1 (c). Subsequently, for each entity, the obtained molecules [1] are processed to construct the triadic formal context. As shown in Fig. 1 (d), time-centric triadic concepts and location-centric triadic concepts can be obtained utilizing TFCA. Finally, a fusion strategy is adopted for integrating the obtained two groups of triadic concepts into the entity spatial-temporal evolution summarization.

3.2 The Approach of the Temporal and Spatial Entity Summarization in FKG

This section details the proposed summarization approach. As shown in Fig. 1 (d), time-centric triadic concepts (i.e., predicate set, object set, and time set) and location-centric triadic concepts (i.e., predicate set, object set, and location set) are generated using the TRIAS algorithm. The obtained summarization of the entity not only contains all the information, but also is compact signifying that each property of the entity is distinct.

3.3 Fusion Strategy

The main idea of the fusion strategy is to merge the time and location labels with the same pairs of predicate and object. The fusion strategy has following merits: It produces an entity summarization in FKGs that contains all evolution information with time and location changes. It ensures the compactness of entity, which means that each property of the summary results occurs only once. If the pairs of predicate and object are same, we merge the corresponding time and location labels. We add the unique pairs of predicate and object to the final summarization results. The time-centric triadic concepts and location-centric triadic concepts can be integrated a compact triadic concepts.

4 CASE STUDY

Inspired by the literature [12], we develop various template-based queries for evaluating the feasibility and effectiveness of our proposed approach. For the single entity, the temporal-spatial query and a refined template query are devised, which can dramatically reduce the query cost in massive descriptions of the entity. The template-based queries can be divided into two categories: basic template query and refined template query. The basic template query utilizes the temporal and spatial constraints to refine the summary results. The refined template query adopts specific triple patterns to restrict the resources for obtaining information that user need.

Example 1. *Fig. 2 describes the information of a person across diverse versions of a FKG as the time evolves. In different time stamps, some properties of the person may change and new properties may be added in different time stamps.*

Example 2. *Fig. 3 depicts the same person in diverse versions of a FKG as the location evolves. The person has various properties in different locations.*

For Fig. 2 and Fig. 3, we can define the threshold value for the fuzzy value, which can convert the FKGs into ordinary KGs. For example, when the threshold value is 0.6, we can regard the value that greater than 0.6 as "1" and other values as "0". Then, we construct the triadic formal context of the entity temporal evolution, as shown in TABLE I. For the sake of simplicity, we only list the formal context of the entity in version 2015 and 2016, and other versions of the entity is constructed in the same way. Afterwards, time-centric

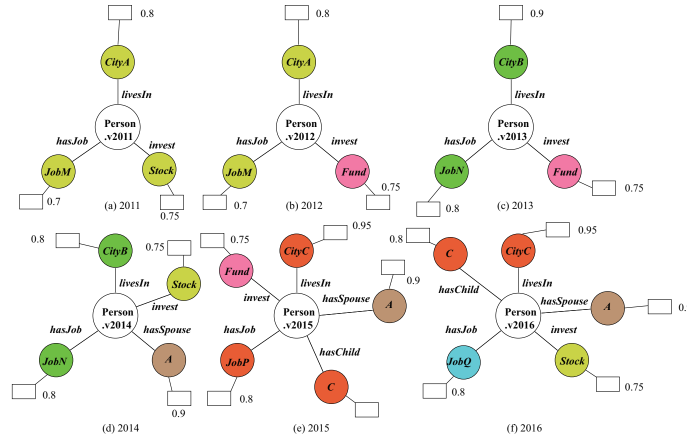


Figure 2: Visual representation of the information of a person obtained from six yearly FKGs.

Table 1: The constructed triadic formal context of the entity in the versions of 2015 and 2016.

	t2015											t2016										
	JobM	*CityA*	*Stock*	*Fund*	*JobN*	*CityB*	*A*	*JobP*	*CityC*	*C*	*JobQ*	*JobM*	*CityA*	*Stock*	*Fund*	*JobN*	*CityB*	*A*	*JobP*	*CityC*	*C*	*JobQ*
hasJob	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1
livesIn	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0
invest	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
hasSpouse	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
hasChild	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0

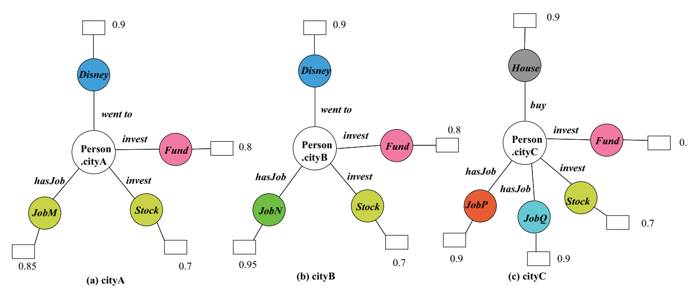
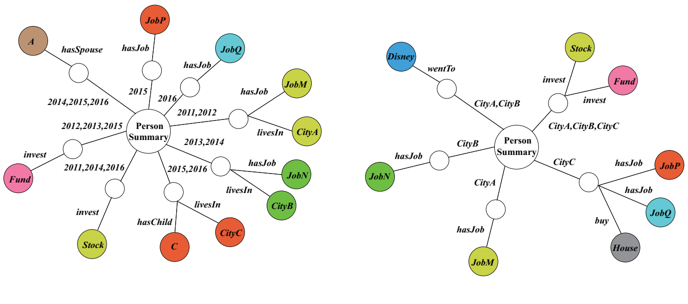


Figure 3: Visual representation of the information of a person obtained from three locational FKGs.

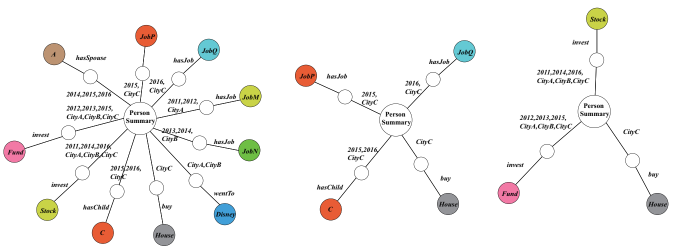
triadic concepts are generated by the TRIAS algorithm [6] according to the obtained triadic formal context, and the results are shown in Fig. 4 (a). Similarly, the results of spatial entity summarization and the corresponding location-centric triadic concepts are shown in Fig. 4 (b). Using the proposed fusion strategy, we can obtain the final triadic concepts. Fig. 5 (a) illustrates the final result of entity spatial-temporal evolution summarization using the fusion strategy.

SPARQL query is the standard query language for retrieving RDF data graph. Thus, we can adopt the basic template query [12] to further reduce the size of result. For example, we can select the person who *livesIn CityC* between 2014 and 2016. For the given entity, the query can be designed as shown in $Q1$. Then the compact result can be obtain as shown in Fig. 5 (b), from which we can see that the size of



(a) Entity Temporal Summarization Representation (b) Entity Spatial Summarization Representation

Figure 4: The entity temporal summarization and spatial summarization.



(a) (b) (c)

Figure 5: The query result with the basic template query.

result is notably reduced. $Q1 : SELECT \quad ?x$
$WHERE \{$
$?x \quad rdf : type \quad dbo : Person.$
$?x \quad ?p \quad ?y. \quad ?x \quad dbo : livesIn \quad db : CityC.$
$?x \quad dbo : version \quad ?version.$
$FILTER(?version \geq 2014).$
$FILTER(?version \leq 2016).$
$\}$

Further, we can apply more specific constraints for obtaining more meaningful information. For instance, we can utilize the $dbo : invest$ and $dbo : buy$ to restrict the size of result. $Q2$ depicts the refined template query and the summary result is shown in Fig. 5 (c).

$Q2 : SELECT \quad ?x$
$WHERE \{$
$?x \quad rdf : type \quad dbo : Person. \quad ?x \quad dbo : invest \quad ?z.$
$?x \quad rdf : type \quad dbo : Person. \quad ?x \quad dbo : buy \quad ?y.$
$\}$

5 CONCLUSIONS

This paper proposes an entity spatial-temporal summarization approach in fuzzy knowledge graph using TFCA. Firstly, we transform the FKG into the ordinary KG by the user-defined threshold. Then, the ordinary KG can be summarized by the proposed fusion strategy. Finally, various template-based queries are designed for evaluating the performance of the proposed approach. The query results demonstrate that the query results can be refined significantly by using the template-based queries. It is expected that these query results would be useful for developing the question-answering systems and recommendation systems in the future.

ACKNOWLEDGMENTS

This work was funded in part by the National Natural Science Foundation of China (Grant No. 61702317), the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 840922, and the Fundamental Research Funds for the Central Universities (GK202103080). This work reflects only the authors' view and the EU Commission is not responsible for any use that may be made of the information it contains.

REFERENCES

[1] Javier D Fernandez, Alejandro Llaves, and Oscar Corcho. 2014. Efficient RDF Interchange (ERI) Format for RDF Data Streams. (2014), 244–259.

[2] Anna Formica. 2012. Semantic Web search based on rough sets and Fuzzy Formal Concept Analysis. *Knowledge Based Systems* 26 (2012), 40–47.

[3] Bernhard Ganter. 1999. *Formal Concept Analysis: Mathematical Foundations*. Springer-Verlag New York, Inc.

[4] Kalpa Gunaratna, Krishnaprasad Thirunarayan, and Amit P Sheth. 2015. FACES: diversity-aware entity summarization using incremental hierarchical conceptual clustering. (2015), 116–122.

[5] Fuzhen He, Zhixu Li, Yang Qiang, An Liu, Guanfeng Liu, Pengpeng Zhao, Lei Zhao, Min Zhang, and Zhigang Chen. 2019. Unsupervised Entity Alignment Using Attribute Triples and Relation Triples. (2019), 367–382.

[6] Robert Jaschke, Andreas Hotho, Christoph Schmitz, Bernhard Ganter, and Gerd Stumme. 2006. TRIAS–An Algorithm for Mining Iceberg Tri-Lattices. (2006), 907–911.

[7] Eun-kyung Kim and Key-Sun Choi. 2018. Entity Summarization Based on Formal Concept Analysis. (2018).

[8] Fritz Lehmann and Rudolf Wille. 1995. A Triadic Approach to Formal Concept Analysis. (1995), 32–43.

[9] Guanfeng Li, Li Yan, and Z M Ma. 2019. An approach for approximate subgraph matching in fuzzy RDF graph. *Fuzzy Sets and Systems* 376 (2019), 106–126.

[10] Denis Lukovnikov, Asja Fischer, Jens Lehmann, and Soren Auer. 2017. Neural Network-based Question Answering over Knowledge Graphs on Word and Character Level. (2017), 1211–1220.

[11] Mayesha Tasnim, Diego Collarana, Damien Graux, Fabrizio Orlandi, and Mariaesther Vidal. 2019. Summarizing Entity Temporal Evolution in Knowledge Graphs. (2019), 961–965.

[12] Xin Wang, Yueqi Xin, and Qiang Xu. 2018. Template-Based SPARQL Query and Visualization on Knowledge Graphs. (2018), 184–200.

[13] Ning Xiong and Peter Funk. 2006. Construction of fuzzy knowledge bases incorporating feature selection. *soft computing* 10, 9 (2006), 796–804.

[14] Erhe Yang, Fei Hao, Yixuan Yang, Carmen De Maio, Aziz Nasridinov, Geyong Min, and Laurence T Yang. 2021. Incremental Entity Summarization with Formal Concept Analysis. *IEEE Transactions on Services Computing* (2021).

[15] Nadezhda Yarushkina, Vadim Moshkin, Aleksey Filippov, and Gleb Guskov. 2018. Developing a Fuzzy Knowledge Base and Filling It with Knowledge Extracted from Various Documents. (2018), 799–810.