Interactive Domain Knowledge Graph Construction for B2B International Trades

Yonghui Xu[1], Gongqi Lin[2], Yuan Miao[2]

Joint SDU-NTU Centre for Artificial Intelligence Research (C-FAIR), Jinan, China
College of Engineering and Science, Victoria University, Melbourne, Australia

xu.yonghui@hotmail.com, gongqi.lin@vu.edu.au, yuan.miao@vu.edu.au

Abstract

Building a cross-border B2B industry knowledge graph can provide the necessary knowledge system for International Core Business Unit (ICBU) of E-commerce companies to move toward intelligent decision making in business and become the necessary foundation for the overall intelligence of ICBU in the future. Knowledge graph and Ontology research are in the same vein, that is, structuring the knowledge of entities, concepts and relationships and providing services. Knowledge graphs in general-purpose domains are large in scale and widely used, such as Wikidata, while knowledge graphs in vertical industry domains are small in scale and restricted in domain, such as commodity knowledge graphs in Taobao scenarios. High-quality industry knowledge graphs usually require the participation of many industry experts. This brings a challenge to the construction of cross-border B2B industry knowledge graphs. To solve this problem, this paper explores interactive domain knowledge graph construction for B2B international trade from the perspective of crowdsourcing.

Keyword: Knowledge Graph, Crowd Sourcing, Domain Knowledge.

I. Introduction

Knowledge graph [1] has been well applied in web search and inference. The developed algorithms are able to construct knowledge graphs reasonably well from general structured or semi structured documents, such as Wikipedia. Similar approach is inefficient while being applied in domain

knowledge graph construction. Domain knowledge [2] often involves specific knowledge which are not widely available in public websites or documents. General knowledge sets such as Wikipedia have been accumulated for a long time and have been well processed. Its structure, format and correctness are of high standard for auto or semi-auto processing. Domain knowledge however can be in a variety forms of documents. Crawling and mining can be very inefficient to create quality knowledge graph.

Secondly, the focus of general knowledge sets is to address ontology issue because many terms can have similar meanings and a term can have many meanings. Domain users, particularly international B2B (Business to Business) domain users, often follow rather specific terms, with limited variety. On the other hand, the domain knowledge sets are highly knowledge intensive and the knowledge could have complex structure. The expectation on knowledge structure and the corresponding inference is much higher than general knowledge sets.

II. Knowledge Graph Construction via Crowd Sourcing

Knowledge graph is gaining significant roles in a wide range of sectors and applications, including decision supports to international trades, automate contract negotiation and review and etc. At the same time, building domain knowledge intensive knowledge graph is exceptionally challenging, in term of the types of knowledge, quality of available online resources, and availability of domain experts. This section discusses solutions to address these challenges, which will significantly support the collection of knowledge from vendors, clients and brokers rather than appointing limited experts.

Given the quality of the domain knowledge documents and difficulties in collection and construction of domain knowledge graph automatically, a practical approach would be via crowd sourcing [3]. By providing a friendly interface, vendors, customers and brokers are able express their knowledge for the construction of high-quality knowledge graph.

It then becomes critically important to motivate vendors, customers and brokers to participate and contribute knowledge. Small incentive which applied in 2C scenarios or normal crowd sourcing

scenarios could be ineffective. B2B customers and vendors are normally not interested in small incentives. What could really work is to make the knowledge graph a very helpful decision support tool. When they interact with the decision support system, e.g. justify why a decision is more preferable than another, the knowledge is collected. Their involvement in others' decision support by offering knowledge and justification could also win them reputation and business.

Decision making involves not only the information (the correlations or general linked information represented by general knowledge graph) but more importantly the causal relationships which are the core to their decision making. Therefore, the knowledge graph needs to be extended to accommodate the causal relationship and the decision-making mechanisms to tradeoff conflicting factors. When the system challenges a user's decision by pointing out the broken logic link, or providing additional knowledge link leading to alternative decision, the user would either change his decision, or provide further justification. This process will be very helpful to users. At the same time, it collects quality domain knowledge for knowledge graph construction.

III. Domain Knowledge Verification and Augmentation

In order to obtain a high-quality knowledge graph, it is necessary to develop methods to support knowledge verification and justification among users or between users and systems, support the merge and inference of knowledge, and support decision making based on the knowledge (graph). To achieve this goal, it is better to build a knowledge graph model for B2B domain knowledge, and through interactive knowledge graph construction approach to collect knowledge via crowd sourcing.

In addition, unlike general knowledge, domain knowledge often involves specific knowledge which are not widely available in public websites or documents. Its structure, format, correctness and quality are far from general knowledge set such like Wikipedia. Crawling and mining approach will not be very effective to create high standard knowledge graph. While International Core Business

Unit (ICBU) facilitates a big number of vendors, clients and brokers who have domain knowledge and can contribute. It is a resource not widely available.

To identify complex structure in the domain knowledge graph, it is necessary to develop augmented knowledge graph for the B2B domain knowledge, which is not available in general knowledge graph models. The focus of general knowledge sets is to address ontology issue because general knowledge is used broadly. Many terms can have similar meanings and a term can have many meanings. Domain users, particularly international B2B domain users, often follow rather specific terms, with limited variety. On the other hand, the domain knowledge sets are highly knowledge intensive and the knowledge could have complex structure. The expectation on knowledge structure and the corresponding inference is much higher than general knowledge sets.

We can develop a decision support system based on the augmented knowledge graph model to provide decision support to B2B users, while motivate users to justify their decisions, effectively collect inputs for knowledge graph construction and extension. It is critically important for ICBU to motivate vendors, customers and brokers to participate and contribute knowledge. Small incentive which applied in 2C scenarios or normal crowd sourcing scenarios will be ineffective. B2B customers and vendors are normally not interested in these small incentives. What could really work is to make the knowledge graph a very helpful decision support tool thus they rely on in their trading negotiation process. It is also a goal of constructing the knowledge graph: to support the automation of international trading process.

IV. Conclusion

This paper discusses how to use crowdsourcing to solve the problem of shortage of experts encountered in the construction of B2B interactive cross-border e-commerce knowledge graphs, but there are still more researchable elements in the construction of knowledge graphs. For example, knowledge complement [4] [5] and knowledge representation [6] are all essential parts in knowledge graph construction. In addition, the reward mechanism [7], anti-cheating mechanism [8], and the

setting of user quality degree for the crowdsourcing model oriented to knowledge graph construction all affect the quality and efficiency of knowledge collection, which still need to be studied.

References

[1] Lisa Ehrlinger, Wolfram Wöß: Towards a Definition of Knowledge Graphs. SEMANTiCS (Posters, Demos, SuCCESS) 2016.

[2] Bilal Abu-Salih: Domain-specific knowledge graphs: A survey. J. Netw. Comput. Appl. 185: 103076 (2021).

[3] Man-Ching Yuen, Irwin King, Kwong-Sak Leung, A Survey of Crowdsourcing Systems. SocialCom/PASSAT 2011, pp.766-773

[4] Xuejiao Zhao, Yong Liu, Yonghui Xu, Yonghua Yang, Xusheng Luo, Chunyan Miao, Heterogeneous star graph attention network for product attributes prediction, Advanced Engineering Informatics, vol.51, 2022, pp.1-12.

[5] Yonghui Xu, Shengjie Sun, Huiguo Zhang, Chang'an Yi, Yuan Miao, Dong Yang, Xiaonan Meng, Yi Hu, Ke Wang, Huaqing Min, Hengjie Song, and Chuanyan Miao. 2021. Time-Aware Graph Embedding: A Temporal Smoothness and Task-Oriented Approach. ACM Trans. Knowl. Discov. Data, vol.16, 3, 2022, pp.1-23.

[6] Y Liu, S Yang, Y Xu, C Miao, M Wu, J Zhang, Contextualized Graph Attention Network for Recommendation with Item Knowledge Graph, IEEE Transactions on Knowledge and Data Engineering, vol.1, 2021, pp.1-12.

[7] Naman Goel, Boi Faltings, Deep Bayesian Trust : A Dominant Strategy and Fair Reward Mechanism for Crowdsourcing. CoRR abs/1804.05560 (2018)

[8] Matthias Hirth, Tobias Hoßfeld, Phuoc Tran-Gia, Cost-Optimal Validation Mechanisms and Cheat-Detection for Crowdsourcing Platforms. IMIS 2011, pp.316-321