Critical Path Identification in Supply Chain Knowledge Graphs with Large Language Models

Yaomengxi Han[1,2], Zifeng Ding[1,3], Yushan Liu[1,3],
Bailan He[1,3], and Volker Tresp(✉)[3]

[1] Siemens AG, Munich, Germany
[2] Technical University of Munich, Munich, Germany
[3] Ludwig Maximilian University of Munich, Munich, Germany

Abstract. In the ever-evolving landscape of global commerce, supply chain management (SCM) has gained increasing significance. An important task in SCM is to find critical supply chain paths for a target company because these paths often represent potential bottlenecks in supply networks and thus could be crucial to risk management. The mainstream solution to this task requires supply chain managers to manually review supply chain data to uncover critical paths, resulting in considerable human labor costs. To better study SCM, recent efforts have been made to construct supply chain knowledge graphs (KGs) that connect supply chain-related data from different sources, facilitating the identification of critical paths through KG reasoning. In this paper, we develop an automated approach for critical path identification (CPI) based on supply chain KGs. We encode supply chain KGs into text and use large language models (LLMs) for CPI. LLMs can not only analyze the topological KG information but also leverage their world knowledge for better path identification. We experiment with two popular LLMs, i.e., GPT-3.5 and GPT-4, and find that they are able to do CPI and meanwhile generate reasonable explanations.

1 Introduction

In today's interconnected global economy, effective supply chain management (SCM) plays a key role in entrepreneurial success. As a crucial task in SCM, critical path identification (CPI) in supply networks has recently gained increasing interest. CPI aims to find the significant supply chain paths related to a certain user-interested company. Each supply chain of length $n-1$ follows the format of $(Company_1 \xrightarrow{supplies\ to} ... \xrightarrow{supplies\ to} Company_n)$, where $Company_n$ is the company of interest. These critical paths often constitute potential bottlenecks for specific products or other vital business operations, highlighting their strategic importance [2]. A major obstacle in solving CPI stems from the lack of transparency in supply networks. Recently, Liu et al. [1] show that companies are usually limited to only knowing their direct (tier-1) suppliers without complete knowledge of further tiers of suppliers. Consequently, they struggle to identify longer critical paths in supply chains. To address this problem, Liu et al. focus

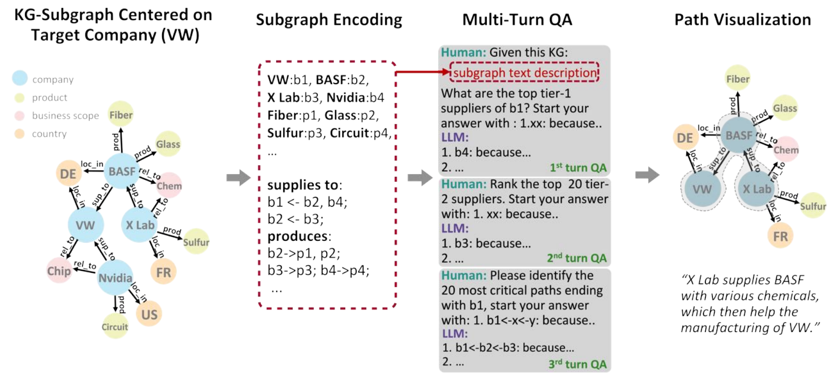

Fig. 1: Framework of approach. Best viewed in color.

on developing transparent supply networks that provide visibility into suppliers up to the third tier while representing supply chains as knowledge graphs (KGs). Despite the introduction of supply chain KGs, [1] has not explored automated approaches to address CPI within supply networks, which heavily rely on manual labor by supply chain managers.

In this study, we present an automated solution for CPI utilizing large language models (LLMs). Given their robust emergent capabilities in diverse downstream tasks without the need for fine-tuning, LLMs serve as a promising tool for this task. Our approach can identify the critical paths given the supply chain KG and any target company. To the best of our knowledge, the method we proposed is the first to use LLMs to address CPI in large supply networks.

2 Approach

The framework of our approach is depicted in Fig. 1. Taking the supply chain KG[4] $\mathcal{G}$ and a target company δ as input, we first extract the relevant subgraph of δ from $\mathcal{G}$ by picking all the KG facts containing δ and all its tier-1 to tier-3 suppliers. Then, the subgraph is encoded into a text description. Based on the description, we initiate a multi-turn question answering (QA) process with an LLM, e.g., GPT-4, in order to step-by-step guide the LLM to provide critical supply chain paths[5] along with corresponding explanations. The criticality of these paths, as well as the consistency between the explanations and the supply chain KG, will then be evaluated by domain experts. Finally, a visualization will be generated for each path for better user understanding. Note that although the LLM-identified paths will be verified by humans, it is much easier than manually identifying critical paths from scratch.

[4] Our work is developed on top of the supply chain KG proposed in [1]. Please refer to [1] for detailed ontology and KG statistics.

[5] In our use case, we only pay attention to the paths of length 2, among 3 entities.

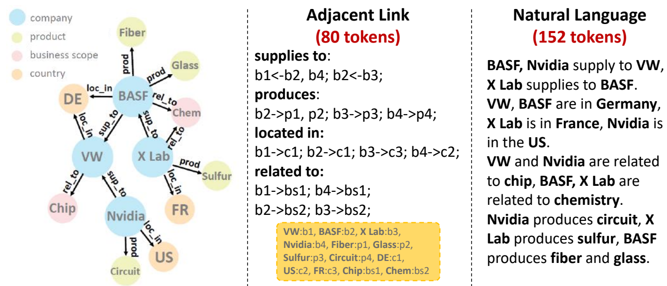

Fig. 2: Two encoding methods. For AL encoding, the contents in the yellow box are not passed to LLMs but just serve as a reference for entity ID mapping.

How to encode a KG subgraph as LLM input? We design two encoding schemes, namely adjacent link (AL) encoding and natural language (NL) encoding. AL encoding (1) maps the names of all entities to distinct IDs, (2) groups the relationships by their types, and (3) translates the relationships into lists of adjacent links. For example, a link $b01 \longleftarrow b03, b05, b11$ under the group **supplies to** indicates that $b03$, $b05$, and $b11$ all supply to $b01$. On the other hand, NL does not anonymize the elements in the supply network but directly outputs a natural language description of it. While NL encoding is more interpretable by both humans and LLMs, it requires more tokens to convey the same information compared with AL. Fig. 2 shows the outcomes of these two schemes, from which we can see this trade-off of token length and interpretability.

Why and how to use multi-turn QA? CPI on a large supply chain KG is a non-trivial task, hence we decompose the complicated task into several sub-tasks. A supply chain path of length 3 consists of the target company, one tier-1 supplier, and one tier-2 supplier. Hence, we start by asking the LLM to give the top 20 significant tier-1 suppliers of the target company, then proceed to ask about the top 20 significant tier-2 suppliers, which are direct suppliers from the LLM-generated tier-1 suppliers. We finally ask LLM to find the 20 most critical paths. We show in experiments that employing task decomposition promotes the LLM's performance in finding more reasonable critical paths.

3 Experiments

To evaluate the effectiveness of our approach, we select three target companies coupled with various sizes of supply chain subgraphs, i.e., **BASF**, **Siemens** and **Henkel** (statistics in Table 1 (left)). For each target company, we run experiments with both GPT-3.5 and GPT-4. We discard the paths returned by the LLMs that do not exist in the supply chain KG and take the rest as the identified critical paths. These paths are then evaluated by domain experts, who would check whether (1) the paths are indeed critical for SCM and (2) the generated

Table 1: Subgraph statistics (left) and experimental results (right). **multi** and **direct** mean with and without multi-turn QA (task decomposition), respectively.

	BASF	Siemens	Henkel
# Entity	301	187	276
# Relation	5	5	5
# KG Fact	1426	1336	903

Model	GPT-3.5				GPT-4			
	NL		AL		NL		AL	
	multi	direct	multi	direct	multi	direct	multi	direct
BASF	0.650	0.455	0.438	0.412	0.958	0.875	0.733	0.600
Siemens	0.714	0.450	0.647	0.550	0.870	0.800	0.706	0.678
Henkel	0.778	0.538	0.650	0.571	0.895	0.786	0.727	0.500

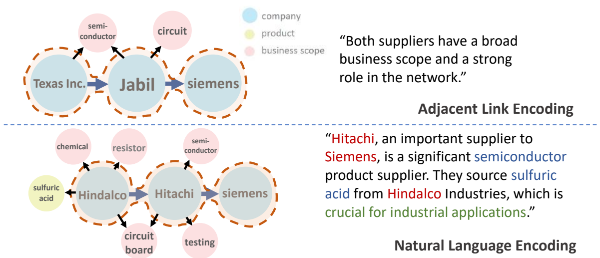

Fig. 3: Two identified critical paths from the Siemens subgraph. Natural language encoding helps the LLM to generate more informative explanations.

explanations are consistent with the supply chain KG. If any of the two requirements is not met, we take the path as incorrect. We let the LLMs return the 20 most critical paths and calculate the accuracy for them. We show the results in Table 1 (right) and demonstrate two correctly identified paths with GPT-4 using different KG encoding strategies in Fig. 3. We observe that (1) LLMs have the ability to automatically do CPI; (2) NL serves as a better encoding strategy since it leverages background knowledge of companies stored in LLMs, making the explanations more reasonable; (3) decomposing CPI greatly helps LLMs to return more accurate critical paths with reasonable explanations since it forces LLMs to pay attention to the critical suppliers that are more likely to exist in critical paths.

4 Conclusion

We propose an automatic approach to encode supply chain KGs and identify critical paths in them with LLMs. Our approach achieves strong performance under the evaluation of domain experts, serving as a new tool that greatly saves human labor in supply chain management.

References

1. Liu, Y., He, B., Hildebrandt, M., Buchner, M., Inzko, D., Wernert, R., Weigel, E., Beyer, D., Berbalk, M., Tresp, V.: A knowledge graph perspective on supply chain resilience. In: D2R2. CEUR Workshop Proceedings, vol. 3401. CEUR-WS.org (2023)
2. Sharma, S.K., Bhat, A.K., Kumar, V., Agarwal, A.: Path analysis model for supply chain risk management. Int. J. Inf. Syst. Supply Chain Manag. **10**(2), 21–41 (2017)