Spatial Knowledge Graph for Analyzing Traffic Accident Data

Seula Park*

* Dept of Civil and Environmental Engineering, Seoul National University; seula90@snu.ac.kr

Abstract. Various attempts are being made to effectively manage and utilize the considerable amount of traffic accident data generated in real-time. A knowledge graph, which organically integrates different contexts based on their relations using graph structures, is a good alternative for managing accident data. This study proposes a traffic accident spatial knowledge graph (TA-SKG) that integrates road-level spatial context to manage detailed traffic accident information. Additionally, a test graph was constructed using emergency response activity data and road networks, and several use-cases for TA-SKG were presented.

Keywords. Traffic accident, Knowledge graph, Graph database

1. Introduction

Road accidents significantly disrupt traffic flow, cause a substantial number of casualties, and damage road infrastructure. To develop prevention strategies, accident patterns can be identified by analyzing accumulated accident records. Consequently, effective management and utilization of accident data collected through various devices is becoming crucial.

Diverse data can be integrated using a knowledge graph, which represents various entities and their relations on a graph structure and provides logical explanations of the relationships between entities (Wang et al. 2019). Analyzing accident data requires comprehensive handling of multiple relevant factors and various contexts; thus, knowledge graphs have been constructed using textual or structured traffic accident data for accident management and analysis (Wang et al. 2019, Liu et al. 2021, Yuhang et al. 2021, Liu & Yang 2022, Zhang et al. 2022).

Published in "Proceedings of the 18th International Conference on Location Based Services (LBS 2023)", edited by Haosheng Huang, Nico Van de Weghe and Georg Gartner, LBS 2023, 20-22 November 2023 Ghent, Belgium.

This contribution underwent single-blind peer review based on the paper.
https://doi.org/10.34726/5748

The locational and topological characteristics of the roads where accidents occur are important factors for accident analysis. Therefore, both road-level spatial context and detailed accident information must be integrated for effective management. However, most previous research has focused on handling the coordinates of the accident locations or spatial characteristics at the urban level.

In this study, a knowledge graph model was proposed to manage traffic accident data by integrating road-level spatial context of accidents. Specifically, a conceptual graph model was designed to include both the characteristic information of accidents and the temporal and spatial knowledge. Additionally, a graph database was constructed for the test area using emergency response activity data and road networks. The suitability of the proposed model was examined through traffic accident analysis and avoidance routing tests using the generated database.

2. Knowledge Graph Model for Traffic Accidents

Entities are defined based on various aspects of traffic accident contexts (Table 1). Specifically, the proposed model represents the accident attributes, temporal and environmental characteristics, and locational and topological features of the occurrence points.

Entity	Description	Attribute
ACCIDENT	Traffic accident	id, latitude, longitude
COUNTRY, CITY, ZONE	Administrative districts where the accident occurred	name
PLACE	The place category where the accident occurred, such as highway, residential road, etc.	place type
FSTATION	The fire station that handled the accident	station, center
CAUSE	The cause of accident	description
TYPE	The type of accident, e.g., Car-to-Car, Car-to-Person, etc.	type
REPORT	The accident report	report number
YEAR, MONTH, DAY, WEEK, HOUR	The time when the accident was reported	time
JUNCTION	The junction of the road	latitude, longitude

Table 1. Core entities of the TA-SKG.

Reported accidents are represented by the ACCIDENT node, whose attributes are the latitude and longitude coordinates of the occurrence point. ACCIDENT serves as a core entity and is connected to other entities through the appropriate relationships. By categorically modeling accident-time information, the accident information can be easily classified temporally, such as by year or day of the week. Based on the accident type, the ACCIDENT node is connected to the corresponding TYPE node through the TYPED_AS relationship. Similarly, the ACCIDENT node is connected to the appropriate CAUSE node through the CAUSED_BY relationship based on the accident cause and to the PLACE node through the OCCURRED_AT relationship based on the occurrence point category. Additionally, the CONDITION entity includes environmental factors at the accident time, such as temperature, humidity, and visibility, and is connected to the ACCIDENT node through the AFFECTED_BY relationship.

In addition to connecting the ACCIDENT nodes to the occurrence place information after decomposing the place into country, city, and zone, the road-level spatial context of the detailed occurrence point is reflected in the model by combining the road network with accident information. The node-edge of the network corresponds to JUNCTION-CONNECT in TA-SKG, representing the topology of roads. The intrinsic attributes of road segments, such as road hierarchy, maximum permissible speed, and number of lanes, can be stored as properties of the CONNECT relationship. The ACCIDENT node is connected through a LOCATED_AT relationship to the two JUNCTION nodes of the road where the accident occurred. Accounting for the road direction, the connection is created from the start node of the road and passes through the ACCIDENT node to the end node (Figure 1). If an accident occurs at a junction, the node is bidirectionally connected only to that single JUNCTION node through an OCCURRED_AT relationship. Consequently, various spatial analyses of accident points can be performed using integrated spatial context. For example, the coordinates stored within the JUNCTION nodes and the network distance stored within the LOCATED_AT relationship can be used for accident-avoidance routing.

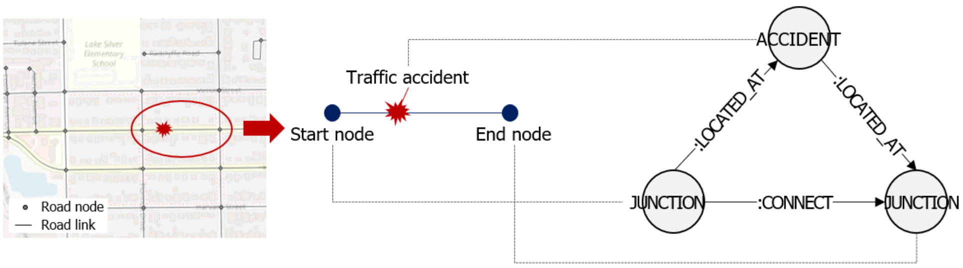

Figure 1. Relationship between accident and junction nodes.

3. Applications

This section introduces the use-cases of TA-SKG. Here, the TA-SKG was constructed as neo4j graph database using emergency response activity data comprising information regarding accidents that occurred in Seoul, South Korea, in 2022. Simultaneously, a spatial graph was constructed for Seoul based on the node-edge connectivity information obtained from the OpenStreetMap (OSM) network.

First, the graph enables easy information retrieval of specific accidents and generates accident descriptions. As shown in Figure 2, relevant information can be found by exploring nodes connected through various relationships related to a particular accident. Moreover, by employing relationship types and node labels, the following descriptions can be readily generated:

- At 1 AM on Saturday, February 26, 2022, a traffic accident was reported. The accident occurred on the highway in Guemhodong-4Ga, Seoul. It was a Car-to-Facility accident and was handled by the Dongdaemun fire station.

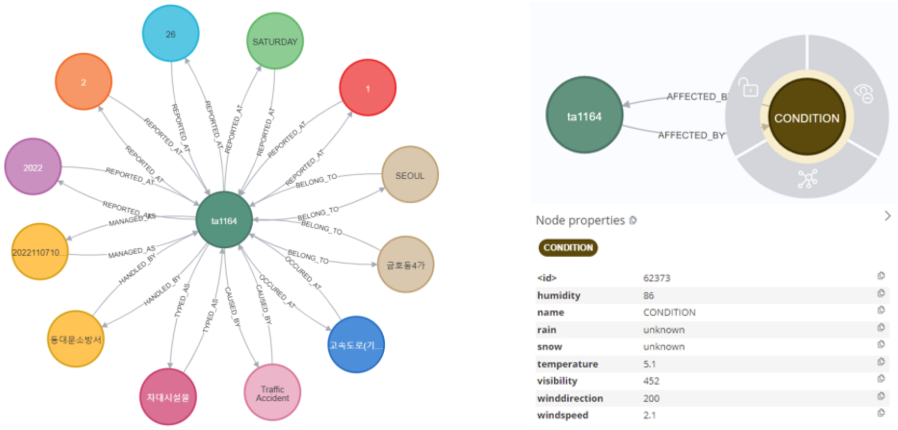

Figure 2. Example of an accident retrieval result.

(a) (b)

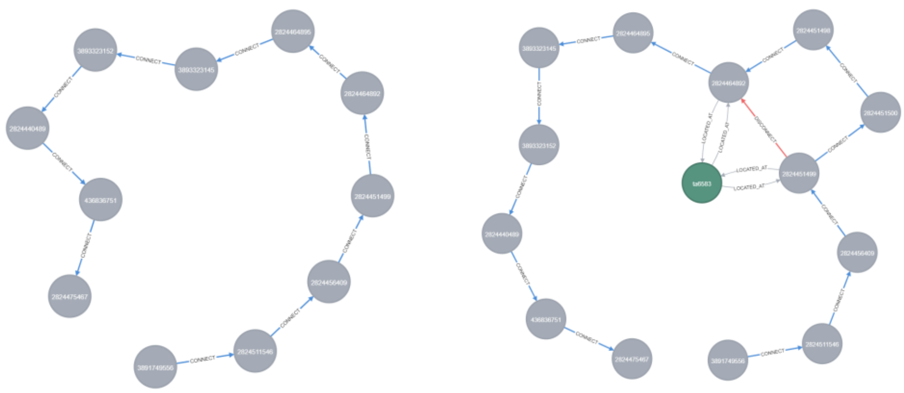

Figure 3. Result of accident-avoidance routing: (a) without and (b) with accident.

Figure 3 shows the result of avoidance routing using TA-SKG after setting an arbitrary path through the road where the accident occurred. It can be seen that the route bypassing the link with an accident is rederived after changing the relationship type (CONNECT to DISCONNECT) between the two nodes (IDs: 2824451499-2824464892) where the accident occurred (Figure 3b) and re-projecting the graph.

Accident information can be easily classified using various node labels, allowing accident pattern identification. Furthermore, temporal analysis of accidents, such as identifying the time of day with the most occurrences, can be performed through time-related entities and relationships.

4. Conclusion

This study proposed the TA-SKG, which combines traffic accident information with road-level spatial contexts. In subsequent research, the data model will be modified to allow integration with existing ontologies of relevant knowledge graphs. Furthermore, the TA-SKG will be expanded using unstructured data sources such as SNS data or accident report documents.

Acknowledgment

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. RS-2023-00209619)

References

Liu C, Yang S (2022) Using text mining to establish knowledge graph from accident/incident reports in risk assessment. Expert Systems with Applications 207

Liu J, Schmid F, Li K, Zheng, W (2021) A knowledge graph-based approach for exploring railway operational accidents. Reliability Engineering & System Safety 207

Wang X, Wang J, Han J (2019) Knowledge graph construction for railway electrical accident analysis. In IEEE 2019 International Conference on Machine Learning, Big Data and Business Intelligence (MLBDBI):214–219

Wang H, Wang Z, Hu S, Xu X, Chen S, Tu Z (2019) DUSKG: A fine-grained knowledge graph for effective personalized service recommendation. Future Generation Computer Systems 100:600–617

Yuhang CH, Jianqin ZH, Jiangchuan LI, An ZH (2021) Visual Mining and Analysis Method of Text Data in Traffic Accident. Journal of Computer Engineering & Applications 57(21):116–122

Zhang L, Zhang M, Tang J, Ma J, Duan X, Sun J, Hu X, Xu S (2022) Analysis of traffic accident based on knowledge graph. Journal of Advanced Transportation 2022