Review Article

# Title: Improving transparency and security of supply chain using Blockchain and Knowledge Graphs

Sahil Arora
19BCE1366
Capstone 2022-2023

Guide - Dr.D.Jeya Mala

# Introduction

The current supply chain system is riddled with disruptions and inefficiencies. The Covid-19 outbreak demonstrated to us how simple it is to completely disrupt the supply chain, leading to the downfall of entire economies. With the help of knowledge graphs and blockchain, we hope to:

- Increase traceability of goods in supply chain to ensure corporate standards are met
- Lower losses from counterfeit/gray market trading
- Improve visibility and compliance over outsourced contract manufacturing
- Reduce paperwork and administrative hurdles

A shipping company transporting goods across multiple transit points would require active tracking, review and approval causing lots of paperworks in the process. This creates opportunity for fraud at multiple points. By executing smart contracts at each stage to automate the management of products and freight, we hope to reduce this process through blockchain and the ledger. Contract manufacturing that is outsourced might be more tightly managed by businesses. Blockchain possibly reduces communication or data transfer errors by giving all participants in a given supply chain access to the same information. It is possible to supply goods and services more quickly by spending less time confirming data and more time enhancing quality. Blockchain and the data stored in it exist only in the digital realm. Interfacing the data tracked in the blocks to the physical goods being traced requires other technologies. By implementing blockchain technologies, stakeholders can

better manage the supply chain by keeping track of information such as price, date, location, quality, certification, and other pertinent details. The availability of this data within blockchain can improve visibility and compliance over outsourced contract manufacturing, increase traceability of the material supply chain, reduce losses from gray market and counterfeit products, and possibly strengthen an organization's position as a pioneer in ethical manufacturing. Businesses can improve their supply chain management through more transparent and accurate end-to-end tracking. We hope to deliver the optimal routes based on a variety of parameters, improving logistical efficiency, using Neo4j and knowledge graphs.

We strive to create a novel solution that would make logistics secure and more transparent using blockchain and distributed ledger.

# Background Work

The first step required for creating a supply chain network was to find a dataset which can be used to create a network of nodes which have various connections. These nodes should have properties which have real world significance and also a real world use case. Such a dataset was provided by the Neo4j GDS core service. This dataset contains the following nodes -
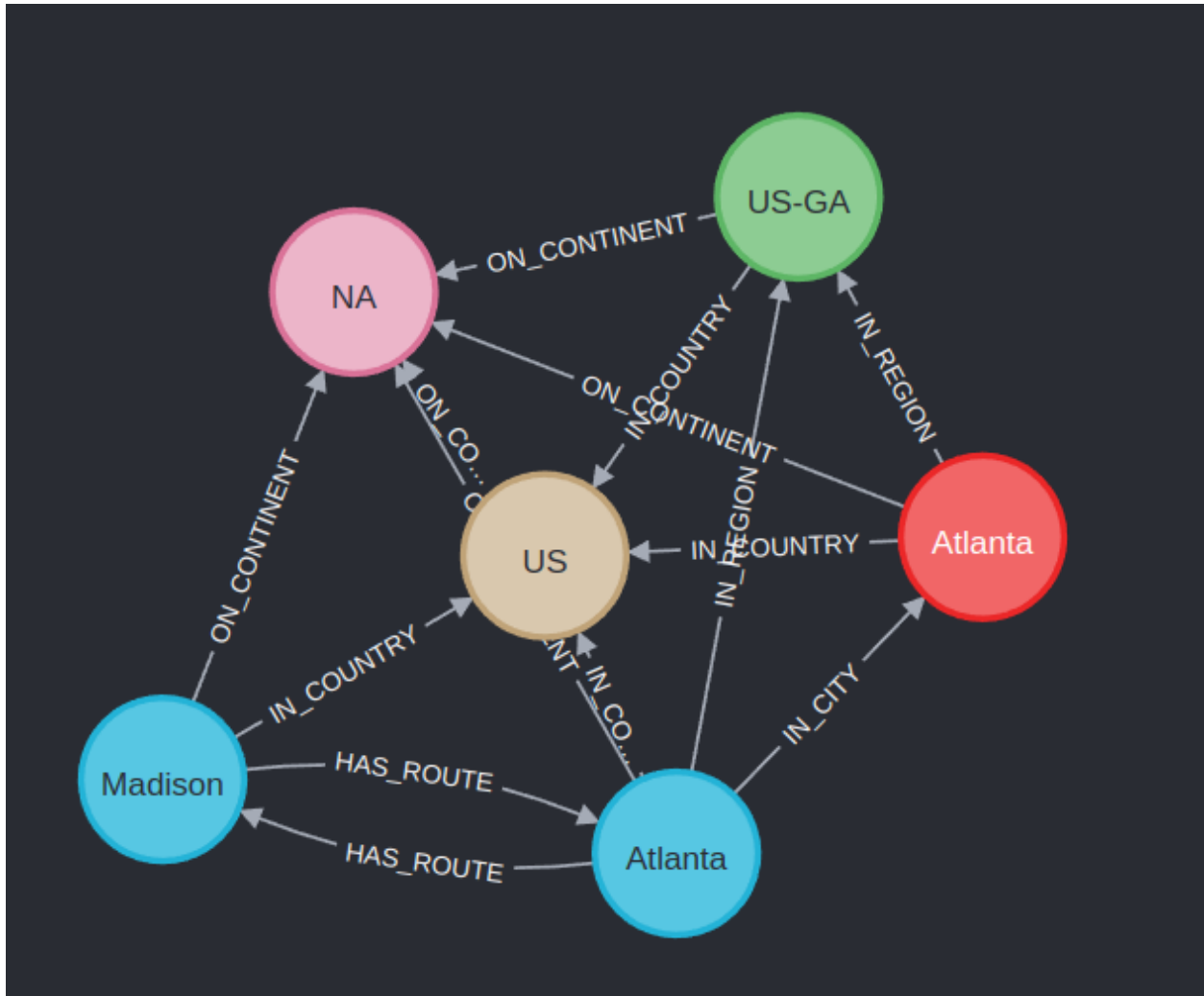
- Continent - 6
- Airport - 3503
- Region - 1527
- Country - 232
- City - 3359

Furthermore the total number of relations connecting these nodes were over 73 thousand. In those 73 thousand relations, 9355 relations signified the routes which are connected.

This dataset had to be used to form a Neo4j graph. The structure (Schema) of the Neo4j graph is given below -
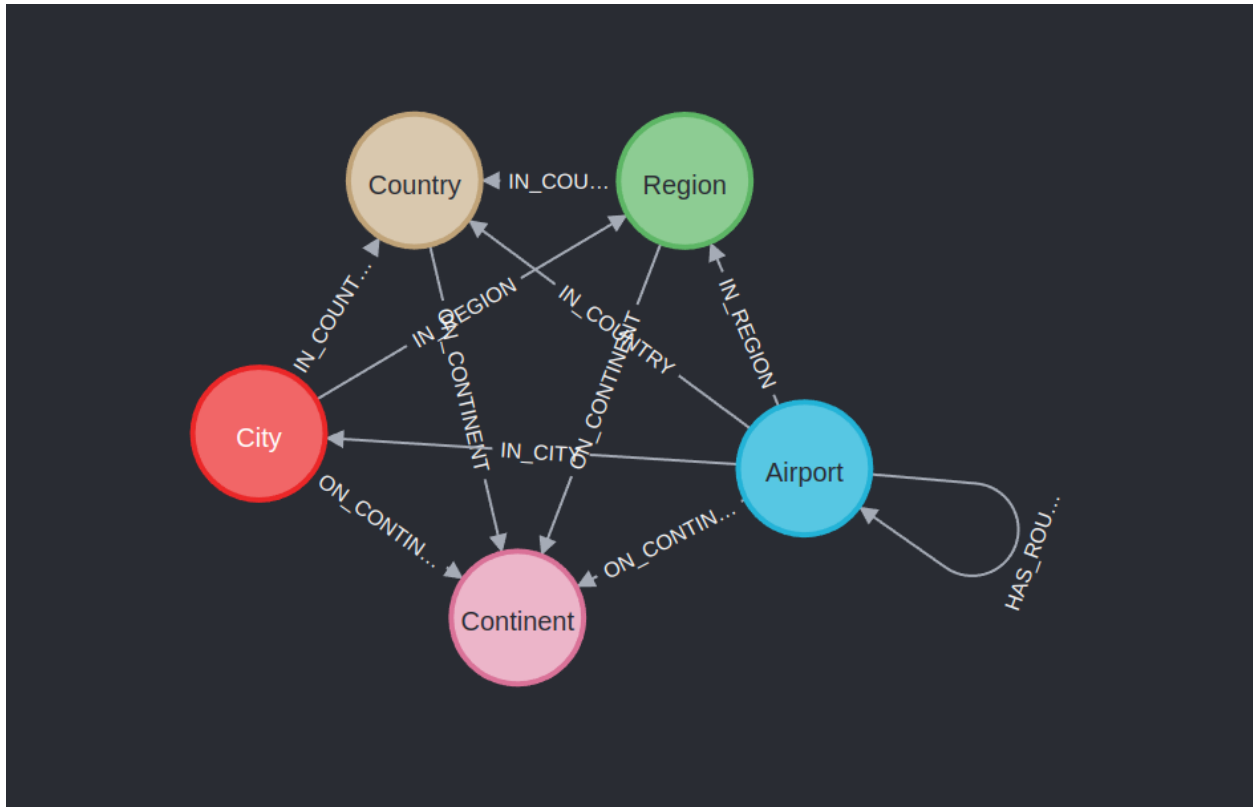
Neo4j Graph Schema

Example Graph with ()->() set to 5

Creating a storage system using blockchain is the next step. Smart Contracts are used to create transactions which can be used to store information in the blockchain. While there are other ways to store information in blockchain, such as using Web3 Storage, using a rudimentary system with less scope of failure is reduced. Solidity was used to write the smart contract.

Solidity is a high level language object oriented language which is used to create and deploy smart contracts. Smart contracts are programs which govern the behavior of accounts within the Ethereum state. Data Structures need to be defined within the smart contract to store data as a blockchain object.

# Literature Review

1. **When Blockchain Meets Supply Chain: A Systematic Literature Review on Current Development and Potential Applications** - The goal of this study, which analyzes 106 review papers, is to examine the existing state, possible uses, and future development of blockchain technologies. It gives us a better understanding of the people working to develop blockchain technology and make them resilient to the future.

2. **Digital Supply Chain Transformation toward Blockchain Integration** - The purpose of this study is to comprehend how blockchain integration functions in business-to-business interactions where service deliveries need to be tracked to give supply chain insight. They look into the specifications and capabilities of such a supply chain integration.

3. **Blockchain technology: implications for operations and supply chain management** - The goal of this article is to promote the study of blockchain technology from an operations and supply chain management (OSCM) perspective, identifying promising application areas and sketching out a methodological approach for the future.

4. **Graph Algorithms - Practical Examples in Apache Spark and Neo4j** - This is a useful book on graph algorithms and how to use the connections and relationships in the data to create clever answers to problems in the real world. This book provides more details on community detection using machine learning-based predictions and Neo4j.

5. **Learning Neo4j 3.x -  Effective Data Modeling, Performance Tuning and Data Visualization Techniques in Neo4**j - In order to improve performance tuning, this book teaches readers how to use cypher-based queries for data modeling and data visualization. This book will teach you the benefits of using graph databases as well as data modeling approaches for graph databases, starting with a quick introduction to graph theory.

6. **Neo4j in Action** - This book talks about how to implement neo4j in the real world. This book is a complete manual for Neo4j graph data design, implementation, and querying. The full potential of native Java APIs for manipulating and accessing graph data is explored in this book. It also discusses Cypher, the graph query language for Neo4j.

7. **Representing temporal knowledge in conceptual graphs** - This research was conducted as part of the lengthy GENTEXT (Generation of Text) project, which aims to create a system that can produce writings in the French language starting with knowledge structures expressed in a manner similar to Sowa's conceptual graphs.

8. **Situating graphs as workplace knowledge** - We look into how graphs are used and understood in the setting of a sizable manufacturing facility. This essay investigates the transparency metaphor's theoretical boundaries. They contend

that a more comprehensive account describing a range of transparency is required, and we aim to do so by utilizing some multimodal approach components that complement a solely activity-theoretical view.

9. **Visualization with Neo4j. Empowering Leaders to operationalize Lean Structural Networks** - This paper works on the organizational hierarchy knowledge graph generation, to create better and more efficient operational structural networks. This study presents a self-configurable, intuitive tool for HK visualization built on top of a Neo4j graph database.

10. **Supply chain network optimization** -  A review of classification,  models, solution techniques and future research - In this study, they take a thorough look at earlier research in the area of SCN optimization strategies and classify supply chain operations according to its relative significance.

11. **Mutually Coordinated Visualization of Product and Supply Chain Metadata for Sustainable Design** - This paper provides a revolutionary visualization framework in this research, in the context of product and supply chain metadata decisions pertaining to redesign. The strategy is based on the concept of overlaying metadata relating to products onto interactive supply chain graph representations and the architecture of the related products.

12. **An Alternative Paradigm for Developing and Pricing Storage on Smart Contract Platforms** - The alternative paradigm for creating smart contracts that is presented in this work, allows for the verification of application data that is stored on and accessed from an external, potentially unreliable storage network and has a continuous state.

13. **Information technology of Blockchain: database, smart contracts, architecture** - The topic of the article is blockchain technology. It is discussed whether databases may be created utilizing blockchain technology. The straightforward process for writing data to the blockchain is accomplished. The primary application domains for technology are listed. Blockchain smart contracts are displayed. This technology's potential is discussed.

14. **Application Response Time Comparison Between Ethereum Smart Contracts And SQLite Database** - In order to assess the viability of using Ethereum Smart Contracts in place of relational databases for developing web applications, this work measures the response time for storing and retrieving data using an Ethereum Smart Contract application. The response time as the user experience for database choice decisions in the construction of future applications was compared between the Ethereum Smart Contract and SQLite.

15. **Smart Contracts for Cooperative Governance** - The research subject of how and to what extent blockchain technology and smart contracts can be used in cooperative governance is covered in this essay, along with how it supports cooperative business structures for sustainable entrepreneurship. Expert

interviews, surveys, and case studies in the context of GTC-related EU initiatives served as the foundation for this study.

16. **Blockchain and supply chain management integration: a systematic review of the literature -** The goal of this study, which analyzes 106 review papers, is to examine the existing state, possible uses, and future development of blockchain technologies. It gives us a better understanding of the people working to develop blockchain technology and make them resilient to the future.

17. **Impact of COVID-19 on logistics systems and disruptions in the food supply chain -** Every type of unit from various sectors is experiencing COVID-19's effects. A simulation model of the public distribution system (PDS) network is created in this study with three different scenarios to show how the food supply chain can be disrupted. Logistical Impact and re-routing were specifically noted in this study Required additional factors for re-routing protocols and how it is managed. The rerouting currently is linear in nature.

18. **Integrating Machine Learning, Radio Frequency Identification, and Consignment Policy for Reducing Unreliability in Smart Supply Chain Management -** In the context of smart supply chain management, this study suggests a machine learning (ML) approach for on-demand forecasting. In order to reduce overstock or understock, the demand is projected using Long-Short-Term Memory (LSTM) to obtain precise demand information.

19. **Modelling Supply Chain Information Collaboration Empowered with Machine Learning Technique -** This paper talks about collaborating supply chains and model them using SVMs. Handles multiple B2B supply chains well, with multiple conflict handling. unreliable the set of features in social media, social events, and location-based store demographic data

20. **Application of machine learning techniques for supply chain demand forecasting** - In this study, they examine the viability of using advanced machine learning methods, such as support vector machines, recurrent neural networks, and neural networks, to estimate distorted demand at the end of a supply chain.Combination of supervised learning, unsupervised learning and machine learning to find results was effective. Better pattern recognition can be used for route prediction.

21. **Predicting supply chain risks using machine learning - The trade-off between performance and interpretability** - They suggest a paradigm for predicting supply chain risks based on the collaboration of supply chain specialists and data-driven AI. The applicability of the framework is demonstrated through a real-world case study of a multi-tier aerospace manufacturing supply chain affected by the risk of delayed deliveries Extracting knowledge through data-driven techniques and using it to derive managerial insights and influence supply chain decision-making process

22. **The Performance Evaluation and Resilience Analysis of Supply Chain Based on Logistics Network** - This study models and analyzes the supply chain using complex networks as a theoretical research tool, and it develops a supply chain network evaluation system that includes static and dynamic measurement indexes in accordance with the statistical properties of complex networks. This study can support decision makers in a wide range of decisions in the context of supply prepositioning to prepare for and respond to disasters. This paper used a case study approach, which restricts the immediate generalizability of insights.

23. **Using Neo4j for Mining Protein Graphs: A Case Study** - Given a protein structure, the aim of the PPI identification task is to pinpoint the amino acids that are necessary for the structure to bind to other proteins. The use case described in this paper is quite a data intensive one operating over a relatively large graph. Lack of any substantial results, mostly a comparison between databases

24. **Decentid: Decentralized and privacy-preserving identity storage system using smart contracts** - This paper suggests DecentID, an entirely decentralized identity storage solution that doesn't need a centralized trusted third party. As a trust anchor, a public blockchain is employed instead. This method is reusable and reliable, with the use of ethereum smart contracts. Interface is not defined, a random structure is prepared and used for authentication of the user.

25. **Dstore: a distributed cloud storage system based on smart contracts and blockchain** - In this article, they suggest a peer-to-peer networking environment for building the DStore client-side encrypted distributed cloud storage system. Without relying on centralized management by vetted third parties, DStore enables data owners to rent the local idle discs of other peers to store personal data in a distributed manner.Reliable file storage which allows for files to be stored in blocks which AES encrypted. No standardization of the file storage smart contract, which was written in solidity.

26. **Machine Learning Model for Smart Contracts Security Analysis** - They presented a machine learning predictive model in this paper that recognises patterns of security flaws in smart contracts. More than 1000 smart contracts that were confirmed and used on the Ethereum platform were labeled using two static code analyzers that we modified. Useful in finding out the standard issue of smart contracts. This is mostly code analysis, Models are trained to tell if some patterns are visible in the solidity code or not. Cases of false positives are visible.

27. **Route planning on GTFS using Neo4j** - The purpose of this study is to analyze how Neo4j, a database management system, can enable public transportation route planning systems using GTFS (General Transit Feed Specification) inputs. The graph database model is particularly useful when data connectivity of the

data is as important as the data itself. The graph is not densely connected. There could have been a better model for the graph.

28. **Analyzing Open Shortest Path First (OSPF) Networks with Neo4j and Cypher** - This system analyzes OSPF using Neo4j and finding the best route possible. Neo4j effortlessly produces a competent network diagram diagram is missing many details that a network technician would ordinarily need

# Summary

These research articles / papers provide the knowledge base required to create the system which can be used for our given use case. These papers were able to provide insights into how an holistic approach to complete construction of a supply chain would take place. Furthermore gave meaning to the approach of blockchain and when to use it. These papers gave additional advantages of using Neo4j rather than a relational database, as it is helpful in exploring connected relations which would not be possible with the former.

# References

1. S. E. Chang and Y. Chen, "When Blockchain Meets Supply Chain: A Systematic Literature Review on Current Development and Potential Applications," in IEEE Access, vol. 8, pp. 62478-62494, 2020, doi: 10.1109/ACCESS.2020.2983601.

2. Korpela, Kari, Hallikas, Jukka and Dahlberg, Tomi, "Digital Supply Chain Transformation toward Blockchain Integration", in Proceedings of the 50th Hawaii International Conference on System Sciences, doi: 10.24251/HICSS.2017.506

3. Cole, R., Stevenson, M. and Aitken, J. (2019), "Blockchain technology: implications for operations and supply chain management", Supply Chain Management, Vol. 24 No. 4, pp. 469-483. doi: 10.1108/SCM-09-2018-0309

4. Needham, Mark., Hodler, Amy E.. Graph Algorithms: Practical Examples in Apache Spark and Neo4j. China: O'Reilly Media, 2019.

5. Bruggen, Rik Van., Baton, Jerome. Learning Neo4j 3.x: Effective Data Modeling, Performance Tuning and Data Visualization Techniques in Neo4j. United Kingdom: Packt Publishing, 2017.

6. Fox, Dominic., Vukotic, Aleksa., Abedrabbo, Tareq., Watt, Nicki. Neo4j in Action. N.p.: Manning, 2014.

7. Moulin, B., & Côté, D. (1991). *Representing temporal knowledge in conceptual graphs. Knowledge-Based Systems, 4(4), 197–208.* doi:10.1016/0950-7051(91)90049-8

8. Noss, R., Bakker, A., Hoyles, C., & Kent, P. (2007). *Situating graphs as workplace knowledge. Educational Studies in Mathematics, 65(3), 367–384.* doi:10.1007/s10649-006-9058-9

9. Jiménez, P., Diez, J. V., & Ordieres-Mere, J. (2016). HOSHIN KANRI Visualization with Neo4j. Empowering Leaders to Operationalize Lean Structural Networks. Procedia CIRP, 55, 284–289. doi:10.1016/j.procir.2016.08.023

10. Matinrad, N., Roghanian, E & Rasi, Z. (2013). Supply chain network optimization: A review of classification, models, solution techniques and future research.Uncertain Supply Chain Management, 1(1), 1-24.

11. Bernstein, William & Ramanujan, Devarajan & Kulkarni, Devadatta & Tew, Jeffrey & Elmqvist, Niklas & Zhao, Fu & Ramani, Karthik. (2015). Mutually Coordinated Visualization of Product and Supply Chain Metadata for Sustainable Design. Journal of Mechanical Design. 137. 10.1115/1.4031293.

12. Patsonakis, Chris & Roussopoulos, Mema. (2019). An Alternative Paradigm for Developing and Pricing Storage on Smart Contract Platforms.

13. V. Kuchkovskiy and N. Shakhovska, "Information technology of Blockchain: database, smart contracts, architecture," 2019 IEEE 14th International Conference on Computer Sciences and Information Technologies (CSIT), 2019, pp. 55-59, doi: 10.1109/STC-CSIT.2019.8929885.

14. Ruiz, Renato & Yano, Inacio & Cezar, Julio & Vasconcelos, Souza. (2022). APPLICATION RESPONSE TIME COMPARISON BETWEEN ETHEREUM SMART CONTRACT AND SQLITE DATABASE. 10.5748/18CONTECSI/COM/ITM/6787.

15. Prause, Gunnar. (2019). Smart Contracts for Cooperative Governance.

16. Queiroz, M.M., Telles, R., & Bonilla, S.H. (2019). Blockchain and supply chain management integration: a systematic review of the literature. Supply Chain Management: An International Journal.

17. Singh, S., Kumar, R., Panchal, R., & Tiwari, M.K. (2021). Impact of COVID-19 on logistics systems and disruptions in the food supply chain. International Journal of Production Research, 59, 1993 - 2008.

18. Sardar, S.K., Sarkar, B., & Kim, B. (2021). Integrating Machine Learning, Radio Frequency Identification, and Consignment Policy for Reducing Unreliability in Smart Supply Chain Management. Processes.

19. Ali, N., Ahmed, A., Anum, L., Ghazal, T. M., Abbas, S., Khan, M. A., ... & Ahmad, M. (2021). Modelling supply chain information collaboration empowered with machine learning technique.

20. Carbonneau, R., Laframboise, K., & Vahidov, R. (2008). Application of machine learning techniques for supply chain demand forecasting. European Journal of Operational Research, 184(3), 1140-1154.

21. Baryannis, G., Dani, S., & Antoniou, G. (2019). Predicting supply chain risks using machine learning: The trade-off between performance and interpretability. Future Generation Computer Systems, 101, 993-1004.
22. Sun, W., Li, Y., & Shi, L. (2020, July). The Performance Evaluation and Resilience Analysis of Supply Chain Based on Logistics Network. In 2020 39th Chinese Control Conference (CCC) (pp. 5772-5777). IEEE.
23. Hoksza, D., & Jelínek, J. (2015, September). Using Neo4j for mining protein graphs: a case study. In 2015 26th International Workshop on Database and Expert Systems Applications (DEXA) (pp. 230-234). IEEE.
24. Friebe, S., Sobik, I., & Zitterbart, M. (2018, August). Decentid: Decentralized and privacy-preserving identity storage system using smart contracts. In 2018 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/12th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE) (pp. 37-42). IEEE.
25. Xue, J., Xu, C., Zhang, Y., & Bai, L. (2018, November). Dstore: a distributed cloud storage system based on smart contracts and blockchain. In International Conference on Algorithms and Architectures for Parallel Processing (pp. 385-401). Springer, Cham.
26. Momeni, P., Wang, Y., & Samavi, R. (2019, August). Machine learning model for smart contracts security analysis. In 2019 17th International Conference on Privacy, Security and Trust (PST) (pp. 1-6). IEEE.
27. Vágner, A. (2021). Route planning on GTFS using Neo4j. In Annales Mathematicae et Informaticae (Vol. 54, pp. 163-179). Eszterházy Károly Egyetem Líceum Kiadó.
28. Holden, W. J. Analyzing Open Shortest Path First (OSPF) Networks with Neo4j and Cypher.