Modelling of Traffic Flows and Supply Chains Based on Geospatial Knowledge

A Kolesnikov¹*, P Kikin² and E Panidi²

¹Department of Cartography and Geoinformatics, Siberian State University of Geosystems and Technologies, Novosibirsk, Russian Federation
²Department of Cartography and Geoinformatics, Institute of Earth Sciences, Saint Petersburg State University, St. Petersburg, Russian Federation

*Corresponding author email: alexeykw@yandex.ru

Abstract. The field of logistics and transport operates with large amounts of data. The transformation of such arrays into knowledge and processing using machine learning methods will help to find additional reserves for optimizing transport and logistics processes and supply chains. This article analyses the possibilities and prospects for the application of machine learning and geospatial knowledge in the field of logistics and transport using specific examples. The long-term impact of geospatial-based artificial intelligence systems on such processes as procurement, delivery, inventory management, maintenance, customer interaction is considered.

Keywords: Geospatial knowledge, spatial modelling, machine learning, logistics.

1. Introduction

The modern stage of development of society is characterized by an increase in the accumulation and processing and analysis of various information resources. A characteristic feature of this stage is the tendency of an increase in the share of spatial data in the total volume of storage and processing [2]. At the same time, the importance of methods and technologies for obtaining, transferring and using knowledge and geospatial knowledge is increasing. Today, in many cases, geospatial knowledge does not have an established terminology and descriptions of the methods and specifics of their acquisition, transmission and use. The main feature of geospatial knowledge (geo-knowledges) is the trend towards the use of spatially coordinated objects. Based on this, first of all, the problem arises of formulating the concept of geospatial knowledge [4].

The main theoretical and practical basis for the formation of geo-knowledges is provided by geoinformatics / geomatics. But now, even more than before, geoinformatics is associated with related disciplines: geodesy, photogrammetry, cartography, remote sensing, etc. Therefore, one of the important features of geo-knowledges is solving the problems of storing, transferring, transforming, and using knowledge in these sciences, considering the spatial component.

Also, the issues to be formalized include the mechanisms of cataloging and searching for geo-knowledge, as well as the selection of the best metric for assessing and interpreting the result of their use. This study is part of a project for the development of Asian Russia, and specifically for the development of transport infrastructure. It represents the concept of extending a database with geo-knowledge. In this article, we will consider the methodological foundations of the formation and use of geo-knowledge in general, as well as their application for logistics purposes. The main point of the study was the formulation of the concept, definition, and components of geospatial knowledge.
2. Materials and Methods
In addition to the direct formation of geosciences, the construction of a terminological base based on the scientific approaches used, the construction of ontologies and a semantic description is additionally carried out in order to realize the possibility of constructing a catalog. A working example of a geo-knowledge catalog is ESRI's Living Atlas, which contains pre-trained neural networks for semantic segmentation of satellite imagery. But, speaking of geo-knowledges, the existing catalogs are still far enough from catalogs of NLP models (for example, Huggingface).

The basic structure in most studies for representing geo-knowledge is a graph representation together with semantic information, mainly for finding the most similar objects [1,3]. In this case, the description format of one of the semantic markup languages is used, the most popular of which are RDF and JSON-LD.

It is the graph structure that is basic for solving logistics problems. These tasks can be divided into the following blocks:

- navigation. With the active development of satellite and inertial navigation systems, as well as SLAM methods, navigation problems have become even closer to geoinformatics. On the basis of knowledge, data is formed about the spatial position of the object relative to other spatial objects.
- transportation management. Geo-knowledges allows you to optimize transportation, improve traffic safety and predict logistics flows in the event of changes in transport infrastructure or resources at the end nodes.
- situational management. Allows to manage flows in any non-standard situations associated with emergencies, changes in the intensity of traffic flows, etc.

Next, we will consider the general methodology of geo-knowledge and the possible specifics when using them for logistics tasks.

3. Results
The authors propose a representation of geo-knowledge as a set of data \((S)\) (including the necessary transformations), mechanisms for forming conclusions based on them \((P)\) and methods of assessing \((M)\) and interpreting the result \((I)\):

\[
GK = S \cup P \cup M \cup I
\]  \hspace{1cm} (1)

Data schema \((S)\) is a description of data types (subdivided into mandatory and optional) \((S_D)\) and required transformations \((S_T)\). Conversions can include both standard format conversions, units of measure, and feature engineering, dimension change, etc.

\[
S = S_D \cup S_T
\]  \hspace{1cm} (2)

Source data \((D)\) is a collection of spatially coordinated data \((DG)\) and statistics \((DS)\) defined at one or more time points \((1 \ldots n)\). It is understood that the statistics initially do not have coordinate values but can be associated with pre-existing spatial objects. Taking into account the already indicated interdisciplinary connections in geosciences, data can be presented in a variety of ways: vector cartographic data, satellite images, three-dimensional elevation models, point clouds, tabular statistical indicators, query results obtained from the protocols of map services, etc. Therefore, the necessary transformations \((S_T)\) constitute a very important and no less voluminous part of the data schema \((S)\) in order to form a single geospace from the entire available data set.

\[
D = DG_{T_1} \cup DS_{T_1}, DG_{T_2} \cup DS_{T_2}, \ldots DG_{T_n} \cup DS_{T_n},
\]  \hspace{1cm} (3)

The processing component can represent both simple conditions and complex mathematical models (including neural networks). At the same time, they can be ensembled, as in the case of processing data of any other types. Of all the elements of geo-knowledges, this component has the greatest variability, since it can be implemented by dozens of algorithms, methods and technologies. For some standardization, universal model formats can be applied (for example, ONNX, but it does not support all the necessary data preprocessing methods), but if they are applied to logistics, then there may be difficulties when working with graph models.
Since there is no direct data in geo-knowledges, but they come from outside, then for each new dataset, methods of assessing the quality and reliability of the result obtained \((M)\) are needed. Also, in most cases, a set of spatial data is formed at the output, so specialized metrics are required that take into account both spatial and temporal components (for example, spatial autocorrelation, Gini Index, Getis-Ord, and others).

In addition to directly assessing the results of processing, it is important to interpret \((I)\) that it is the process of forming conclusions \((P)\) that occurs. The existing algorithms of artificial intelligence (machine learning) form the problem of the contradiction between accuracy and interpretability. Deep neural networks make it possible to solve more and more complex problems, but it is very difficult to understand how exactly this result came about. In the proposed concept of geoknowledge, it is assumed to use mainly model-independent (agnostic) approaches, implying operating with models presented in the form of a black box, and only their inputs and outputs are used for interpretation.

It should be noted that in the proposed version of geo-knowledge should not have a graphical representation in the form of a map. The map is obtained based on the results of processing using geoknowledge, which are also spatial and statistical data, but with a greater degree of generalization.

\[
GK: D \rightarrow D_{GK}
\]  

If we concretize the scheme of geo-knowledges related to logistics, then its elements should be economically significant objects (as a separate spatial element, or in generalized in the form of economic regions), connecting them with transport infrastructure of all types. Additional elements will be external factors affecting the integrity and sustainability of the transport infrastructure. These elements mean both negative factors that remove or reduce characteristics (for example, the amount of resource extraction, units of produced goods, the road capacity) of economically significant objects and connections between them, and positive ones that increase characteristics that add new transport links.

For the digital twin of the Asian Russia logistics system being developed, the authors propose the following initial components:

- description of the spatial structure based on maps of the territory;
- location of places for intermediate storage and transshipment of goods;
- structure of input and output streams;
- the structure of dynamic logistics objects that are moved and stored;
- the cost of performing operations in the nodes and arcs of the network;
- characteristics of transport channels;
- strategies for supply chain management;
- current and projected demand values;
- probabilities and characteristics of external factors affecting the nodes and arcs of the network;
- duration of technological operations.

For implementation, a multi-model DBMS is used that allows you to combine the formation of links between data and ontologies. The following types of links between ontologies are used.

- taxonomic relationships - for the implementation of the “is a child” relationship.
- compositional links - to implement the relationship “part-whole” in the form of one of the implementations “component-object”, “ingredient-product”, “set element”.
- topological links - for the implementation of spatial relationships.
- relationships of entities with processes - for the implementation of relations of participation of the entity in the process, the execution of the process or its leadership.
- cause-and-effect relationships - for the implementation of relations of the influence of processes on entities, other processes (call, change, or termination).
- temporary connections - for the implementation of connections between entities flowing in time and between time intervals.
4. Conclusion

Thus, geo-knowledge will make it possible to combine the tools of a geographic information system, a traditional spatial database (for storing and most optimal processing of spatial data, which will be worked with using a geographic information system), a graph database (with extensions that allow storing and using semantic networks), technical solutions that implement algorithms and methods for constructing mathematical models obtained by both direct and reverse modeling. This approach will allow realizing both visualization and a comprehensive understanding of the internal processes and flows of the system and combine the two main approaches to the simulation of logistics networks - graphic and mathematical models. It will also allow the use of common modeling approaches - SCOR, EPCs, EPK, EOQ. Also, the developed concept of geodata will allow solving the tasks of strategic (assessment of the functioning of the production system at large time intervals), tactical (assessment of the current state and efficiency of the functioning of the existing production system), operational scheduling (building a production schedule for a short time - from hours to days, current planning, scheduling the loading of technological equipment, dispatching tasks) and simulation (a tool for playing scenarios of scenarios "what-if").

Acknowledgments

The research was prepared within the grant for carrying out major scientific projects in priority areas of scientific and technological development under the subprogram "Fundamental scientific research for the long-term development and ensuring the competitiveness of society and the state" of the state program of the Russian Federation "Scientific and technological development of the Russian Federation," project "Social and economic development of Asian Russia on the basis of synergy of transport accessibility, system knowledge of natural resource potential, and expanding area of interregional interactions", agreement with the Ministry of Science and Higher Education of the Russian Federation № 075-15-2020-804 (internal number 13.1902.21.0016).

References

[1] Ballatore A, Bertolotto M and Wilson D C 2015 A Structural-Lexical Measure of Semantic Similarity for Geo-Knowledge Graphs ISPRS International Journal of Geo-Information vol 4(2) (MDPI AG) pp 471–492 doi:10.3390/ijgi4020471.
[2] Lin H, Lan Y, Chuanbo H and Min C 2018 Prospect of Geo-Knowledge Engineering in the Era of Spatio-Temporal Big Data Geomatics and Information Science of Wuhan University vol 43(12) (Wuhan University) pp 2205–11. doi:10.13203/j.whugis20180318.
[3] Chen J, Shumin D and Huajun C 2017. CrowdGeoKG: Crowdsourced Geo-Knowledge Graph Communications in Computer and Information Science vol 784:165–72 (Springer Verlag) doi:10.1007/978-981-10-7359-5_17.
[4] Tsvetkov V Ia 2016 Geo-knowledge European Journal of Technology and Design vol 13(3) (Academic Publishing House Researcher)
[5] Lobanov A A 2014. Geo-knowledge as a new form of knowledge Prospects of Science and education vol 5 pp 37-44.
[6] Rudskoy A, Ilin I and Prokhoro A 2021 Digital Twins in the Intelligent Transport Systems Transportation Research Procedia vol 54:927–35. (Elsevier B.V.) doi:10.1016/j.trpro.2021.02.152.
[7] Katarzyna T and Kubik A 2020 Economic Aspects of Driving Various Types of Vehicles in Intelligent Urban Transport Systems, Including Car-Sharing Services and Autonomous Vehicles Applied Sciences vol 10(16) (MDPI AG, Switzerland) doi:10.3390/app10165580.
[8] Tripathi P S M, Kumar A and Chandra A 2021 An Overview of Intelligent Transport System (ITS) and Its Applications Journal of Mobile Multimedia vol 17 (River Publishers) doi:10.13052/jmm1550-4646.17134.
[9] Yujun D, Wang J, Rizos C and El-Mowafy A 2021 Vulnerabilities and Integrity of Precise Point Positioning for Intelligent Transport Systems: Overview and Analysis Satellite Navigation vol 2(1) (Springer Science and Business Media LLC) doi:10.1186/s43020-020-00034-8.
[10] Xingjuan C, Zhaoming H and Jinjun C 2020 A many-objective optimization recommendation algorithm based on knowledge mining, Information Sciences vol 537 pp 148-161 https://doi.org/10.1016/j.ins.2020.05.067.