Evaluation of Topology Description Models in Road Network Formats

János Máté Lógó, Árpád Barsi

Department of Photogrammetry and Geoinformatics, Faculty of Civil Engineering, Budapest University of Technology and Economics
Műegyetem rkp. 3, H-1111 Budapest, Hungary
E-mail: logo.janos.mate@emk.bme.hu; barsi.arpad@emk.bme.hu

Abstract: Topology is a particular feature of road networks. Topology refers to properties, such as the connection of roads, but not only through their axes or reference lines, but also at the level of lanes. The map topology of road networks does not currently have procedures that can be expressed in mathematical formulas (or only to a minimal extent), but can only be implemented by algorithms. Our research, therefore, aimed at observing such topological regularities and then developing a method of investigation, by constructing rules and additional algorithms. To test this work, we used synthetic and real data, focusing on the case of map content embodied in four formats. In this paper, we present the test methods for the data models, the results of our test runs and finally, we point out that topological checks are extensively justified to determine the quality of the produced map. In the future we plan to develop an automatic correction mechanism based on this.

Keywords: road network; topology analysis; autonomous vehicle simulation

1 Introduction

One of the most promising developments of our time is the automation of transport. Computers are increasingly being used in the development of self-driving vehicles, through simulation techniques [1] [2]. Simulations need an accurate and detailed representation of reality, which can be provided by specific map content [3-8]. These maps have “grown out” of traditional navigation products, but have some unique features.

Self-driving vehicles are expected to make autonomous decisions on how to control the vehicle as it moves through traffic. By making safe choices, they can select and follow the best and safest route. This requires support from maps in which the true geometric characteristics of the road infrastructure are given. Beyond pure geometry, however, the importance of topology was recognized very early.

Topology is the science of the invariant properties of objects that are preserved during various deformations or transformations. Examples of such properties are the connection of surfaces without gaps or overlaps, the connection of lines and line chains, or the inclusion of points on them.

In this paper, we will discuss various theoretical and practical implementations of map topology, their analysis and verification, the errors that occur, and how to correct them. In the second section of the paper, we describe the theoretical background of map topology. In the third section, we present four realizations found in practice. The fourth section is devoted to the investigation of the road network topology. We close our study with a conclusion.

2 Topology Models

Map topology is most often described using graphs. By definition, a graph G consists of its vertices and the edges that connect them: $G = (V, E)$, where V is the set of vertices (alternative names are nodes, points) and E is the set of edges (or alternatively arcs, links, lines). For the latter, we define the following relation, which makes sense of the edge fit of the vertices:

$$E \subseteq \{(x, y)\} \mid (x, y) \in V^2 \; and \; x \neq y \quad (1)$$

Instead of storing the data in the above-unordered list format, it is often more appropriate to use the matrix notation, which also helps to keep track of the data. The simplest matrix is the adjacency matrix for the vertex-to-vertex collocation, which is n × n for n nodes. Its elements take values according to the following rule:

$$A_{ij} = \begin{cases} 1 & if \; edge \; (V_i, V_j) \\ 0 & if \; no \; edge \end{cases} \quad (2)$$

The second matrix representation is the edge-vertex representation, called also as incidence matrix. By definition, it is the following:

$$B_{ij} = \begin{cases} -1 & SP \\ 1 & EP \\ 0 & otherwise \end{cases} \quad (3)$$

where *SP* is the starting point, *EP* is the endpoint. The edge-edge representation is defined as

$$C_{ij} = \begin{cases} -1 & P \\ 1 & S \\ 0 & otherwise \end{cases} \quad (4)$$

where *P* means, that the predecessor of V_i is V_j, whilst the successor *S* of V_i is the vertex V_j.

The above matrix forms can be interleaved with mathematical expressions, i.e. the edge-vertex matrix can be used to calculate the other two matrices:

$$\mathbf{A} = \mathbf{B}^{\mathbf{T}} \cdot \mathbf{B} \tag{5}$$

$$\mathbf{C} = \mathbf{B} \cdot \mathbf{B}^{\mathbf{T}} \tag{6}$$

This is of paramount importance in storing data and maintaining and monitoring data consistency [9].

The strongest connection between geometry and topology is revealed in topological errors. The main types of errors are shown in Figure 1.

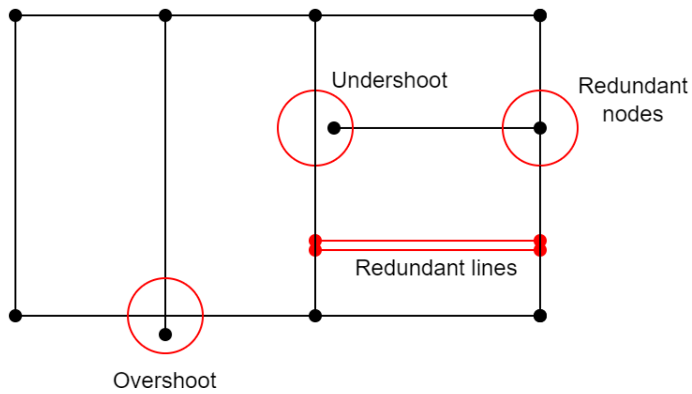


Figure 1
Topological errors

For the road network, directed graphs (digraphs) are preferred, where the edges are oriented.

3 Data Formats for Road Network Storage

Graphs are implemented in concrete formats in order to manage real road networks in geographic information systems (GIS). These realizations consist not only of a data model, i.e. a way of describing the elements of the graph, but also of an application environment. The essential components of the latter are the procedures and operations that can be performed on the data stored in the data model.

Our approach describes four types of solutions. The first is specific to one of the most common GI systems. The developer, ESRI (Environmental Systems Research Institute) **ArcGIS** handles line elements by storing points with their coordinates and then creates an "Arc-Node Topology", the concrete form of which, is the Arc Attribute Table (AAT). This table contains the edge identifiers (ID), start point (From-node), end point (To-node), and may be extended with the identifiers of the polygons on the left and right sides (Lpoly and Rpoly). The storage organized in coverage requires strict rules: 6 points, 15 lines, 10 polygons, and 1 line/polygon basic rules are defined.

Our second example is **OpenStreetMap** (OSM), a map database based on community data collection and mapping. OSM has been created as a collaborative project by Steve Coast in the UK in 2004, initially inspired by the success of Wikipedia. OSM is community-owned, but supported by the OpenStreetMap Foundation. OSM data is stored in a PostgreSQL database with PostGIS extension. For data transfer, dumps are created, which are available in two formats: XML and Protocol Buffer Binary Format (PBF). There are further data providers (e.g., the German Geofabrik), where also ESRI shape (SHP) format data is available [10] [11].

OSM's topological data structure is built up from four core elements (also known as data primitives):

- **Nodes:** Points with a geographic position
- **Ways:** Ordered lists of nodes, representing a polyline, or a polygon
- **Relations:** A relation is a multi-purpose data structure that documents a relationship between two or more data elements (nodes, ways, and/or other relations)
- **Tags:** Key-value data pairs

Topology element nodes are twofold in the strict GIS sense: start or end point – called node, and intermediate point – called vertex.

OpenStreetMap has 8.3 million registered users, contains 7.4 billion nodes, have ~4 million map changes/day from 1.75 million different user contributors. The world's uncompressed XML-format OSM database exceeds the size of 1561.5 GB. (Statistics from [12] and [13] on 2022-04-21). The complete Hungarian road network available from Geofabrik has 7.6 million points (1.7 million nodes and 5.9 million vertices) and 864 thousand polylines.

If two polylines are connected at a T-intersection, in a strict topology, the connection point must be a node, which means, that the previously created line must be broken and a node element must be inserted. In contrast, the OSM topology is more permissive: the endpoint of the line starting at the join must be on the other line, but a "simpler" vertex element is sufficient.

The other very important difference from the strict GIS topology is that edges cannot only be directed. This will result in unidirectional and bidirectional elements being "mixed up" in the database, making it difficult for application developers to implement, for example, route planning (Fig. 2).

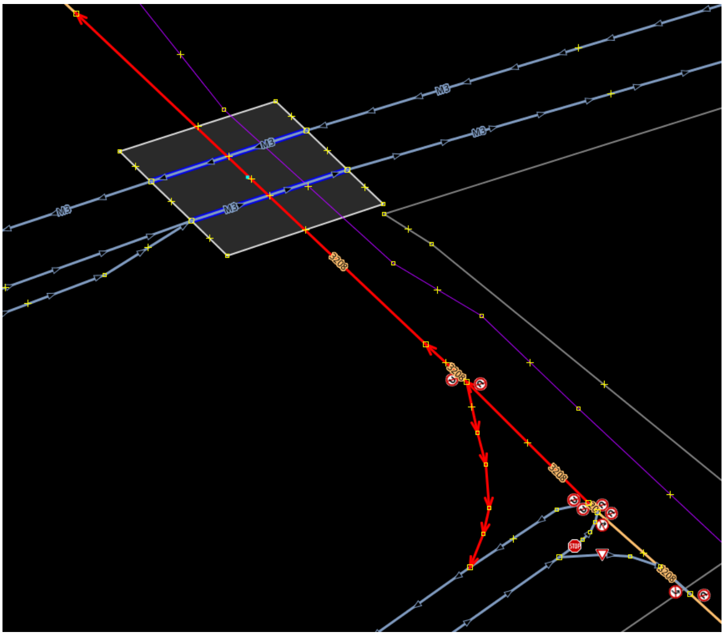


Figure 2

Topology example of a minor road crossing a motorway: both unidirectional and bidirectional edges (see red arrows!) are included in the database (made with the JOSM editor at the junction of the motorway M3 and the rural road 3208)

The third system is the **Navigation Data Standard** (NDS) generally used in the vehicle navigation world [14] [15]. It is a standardized format for automotive-grade navigation databases, jointly developed by automobile manufacturers and suppliers. NDS is also an association registered in Germany with 43 international members of automotive developers, map data providers, and navigation device/application providers.

NDS uses the SQLite Database File Format. An update region represents a geographic area in a database that can be subject to an update. Update regions thus enable incremental and partial updating of defined geographic regions within an NDS database. All navigation data is organized into specific building blocks: 3D objects, Basic map display, Digital terrain model, Full-text search, Junction view, Lane, Name, Natural guidance, Orthoimages, Points of interest, Routing, Shared data, Speech, Traffic information, Volatile data.

In topological terms, NDS is one of the most sophisticated, most mature solutions. It has two data levels: road level and lane level. Fig. 3 illustrates a complex junction with intersecting roads and the corresponding lanes. Roads and lanes are joint with connectors, also numbered and identified in the data model.

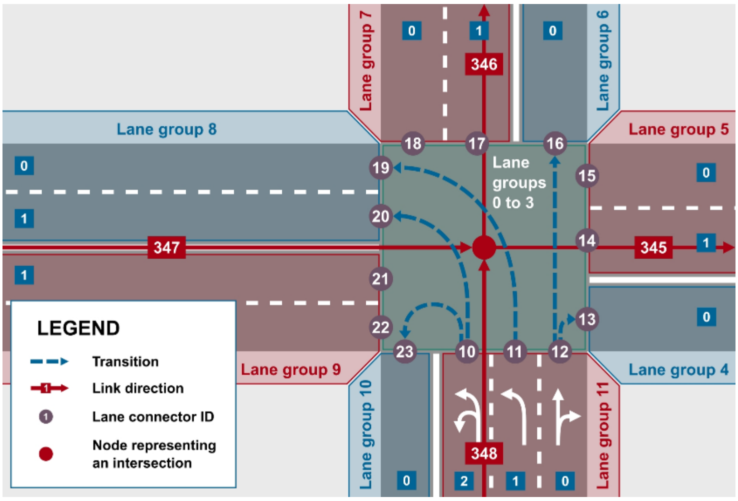


Figure 3

NDS example: A schematic junction with the road and lane level elements [14]

Finally, the last format is **OpenDRIVE**, used in the world of automotive simulators. It was originally a format developed by the German company Vires for various vehicle-oriented simulations to represent real-world map data [3] [4]. It is now standardized; the latest version is 1.7.

OpenDRIVE is used to describe road infrastructure and its environment in 3D. Its XML format provides an inefficient storage for tags encoding many features hierarchically. The format carries the extension XODR.

In terms of topology, the format provides connectivity at the road and lane level using links (Fig. 4). In ambiguous cases (at splitting or merging), a junction must be formed. The road axes are directed, to which the traffic lanes are related. The lanes may point in the same direction as the axis and in the opposite direction (e.g., a bidirectional road). The links of the directional elements can be of predecessor and successor types. From a topological point of view, additional rules can then be introduced, for example, in the case of two connecting roads, taking into account the direction of the axes, the second road can be the successor of the first one, while the second has the first as a predecessor. The possible cases and their implementation are described in detail in the article [9]. It should be emphasized that in OpenDRIVE, we specify the connections not only at the road level but also at the lane level, i.e., the logic of the lane connections must be specified.

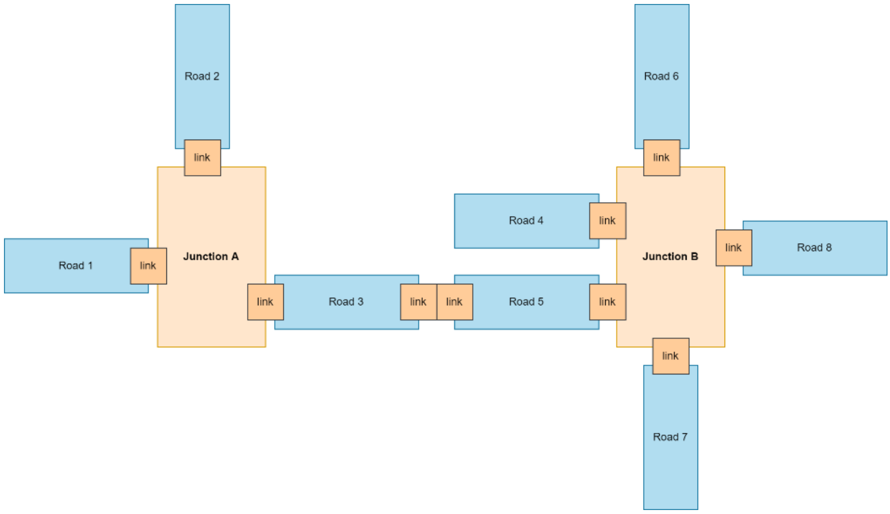


Figure 4
Schema of the OpenDRIVE model with road linkages

4 Evaluation Strategy for Road Topology

The first of the four topological approaches described, the **ESRI** model, is included in this paper because it is considered in GIS as a kind of cord scale; in most cases it is used as a benchmark. Of course, this includes both the topology of coverage and shape formats.

The **OpenStreetMap** model was analyzed in depth with data from 2015. (The analysis software written during the research started a few years ago is outdated due to a change in the data model in 2016; it needs to be rewritten, which is not yet fully completed.) The total road network of the country stored in OSM looked like the following in 2015 (Fig. 5).

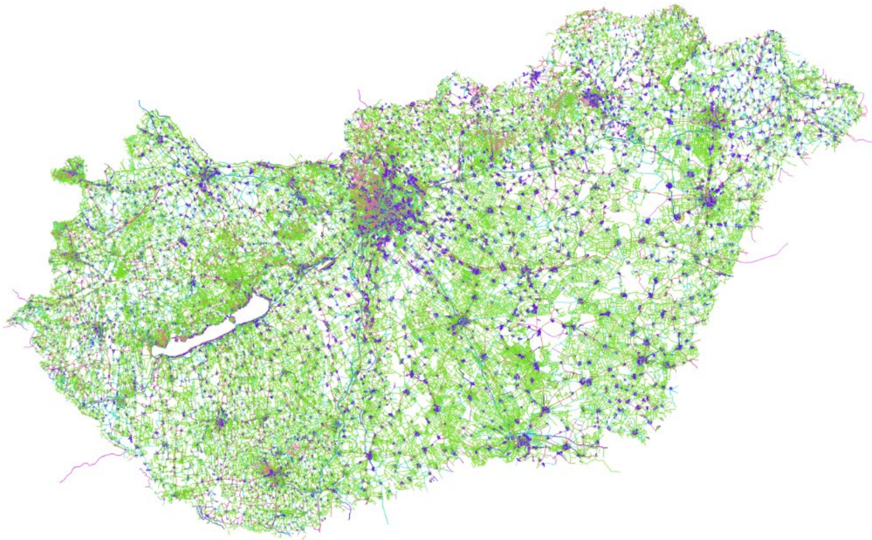


Figure 5
The entire road network of Hungary in 2015 with all 419 076 elements mapped

In the road network, 37 different road categories have been distinguished, but with a huge variation in the number of elements. Fig. 6 displays the frequencies for the eight most important road types. It can be seen that compared to the 140 996 residential types, secondary has only 15 741 items, a difference of almost 9 times (one order of magnitude).

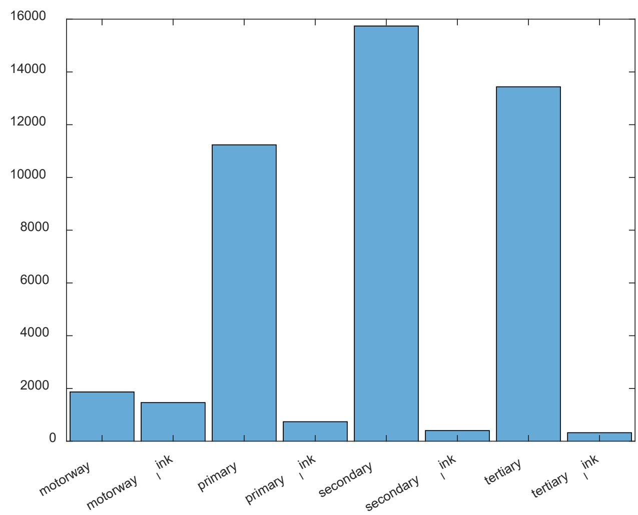


Figure 6
Frequencies of the eight main road categories

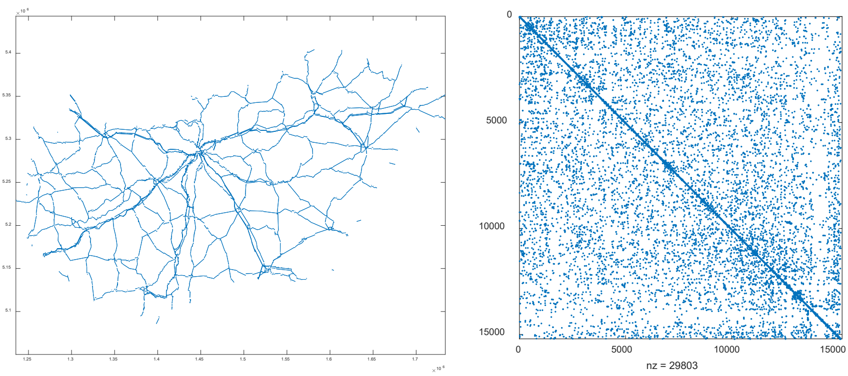
The differences in frequency also draw attention to the fact that the interchange between different road categories and connections is of paramount importance for practical use.

If the network is restricted to the element types motorway, motorway link, primary and primary link, i.e., the main road network, the number of polylines downloaded is 15 307, which in the Geofabrik (format-converted) shapefile represents 142 904 points. A small fraction of this set of points (30 614 points, ~21%) is the node, the rest being the vertex (112 290 points).

For topological verification, we defined a geometric tolerance, so that two points occurring within a radius of 5 m were considered as one. This metric was determined based on the size of the lanes and practical experience. While keeping the tolerance in mind, we rebuilt the topology: we gradually added to the model the points that were found to be different and the edges that matched them. In all cases, we followed the strict topology rules of ESRI.

This analysis resulted in 15 208 independent points, 231 282 978 distinct edges. On an average notebook, the topology build-up was about 29 s.

The connectivity of the edges of the path network, the nonzero elements of its primary adjacency matrix, is shown in Fig. 7. It can be seen that there are no prominent nodes, i.e., the network has a largely uniform topological distribution.



a) the main road network b) primary adjacency matrix

Figure 7
The main road network and the primary adjacency matrix

Very similarly to the adjacency matrix, a distance matrix can be derived, and then by generalizing it to an all-pair type, the distance of each node to each node can be given. Using such an analysis technique, a network reachability analysis can be performed.

To investigate the **NDS** topological model, a sample model of the St. Gellért Square in Budapest was created by orthophoto on-screen digitization (Fig. 8).

The orthophoto was captured in 2019 on behalf of the municipality, from which the mosaics of the working area were merged as Geotiff at a geometric resolution of 0.0625 m in HD72/EOV map projection system. It should also be added that the evaluation was carried out by several independent evaluators. The resulting models were then aggregated and the conflicting positions were discussed and finalized by consensus. The guiding principle for the evaluation was the accurate consideration of traffic rules (e.g., no turning). From the 378 evaluated elements, 9 intersections and 33 lane groups were constructed according to the NDS standard. The generated graph was compared with the data content of the HERE HD map for verification [16]. In the sample area, the complete agreement between the two datasets has proven that the lane design was performed according to the NDS rules. As there are many improvements to the NDS by map data providers, we did not attempt to analyze this model in depth.

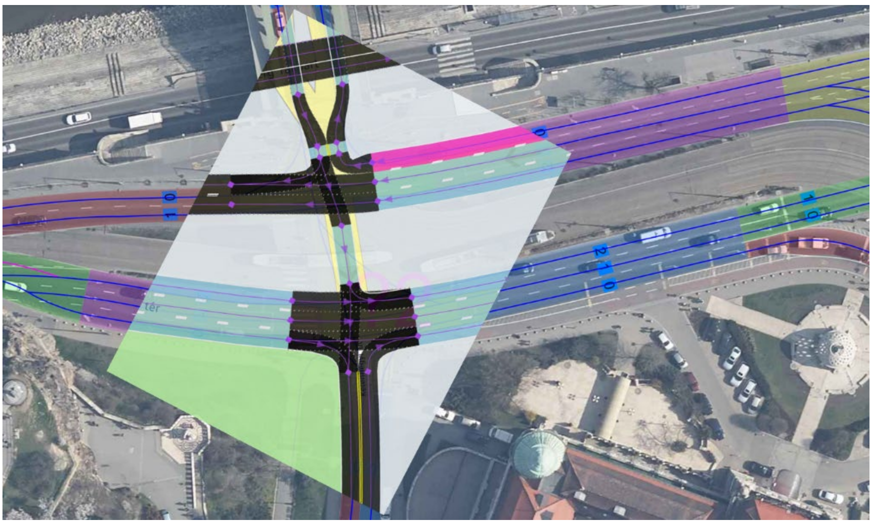


Figure 8
An orthophoto of the sample area with the evaluated NDS elements and a detail of the HERE HD map for comparison

Of the topological models, the **OpenDRIVE** implementation for simulators has received the most attention. This standard is the youngest, has the least experience, and could play a major role in the development of autonomous vehicles. In automotive simulators, standard version 1.4 is the most commonly supported, so we have focused our work on that.

One of the easiest tools to create OpenDRIVE models is the MathWorks RoadRunner software [17]. We used it to create several synthetic test cases and then studied how topology appeared in the models. One of the synthetically designed complex sample spaces (with a roundabout, an X-intersection, and several types of T-intersections) is shown in Fig. 9. It is worth observing how complex the lanes and their relationships can become in turning situations.

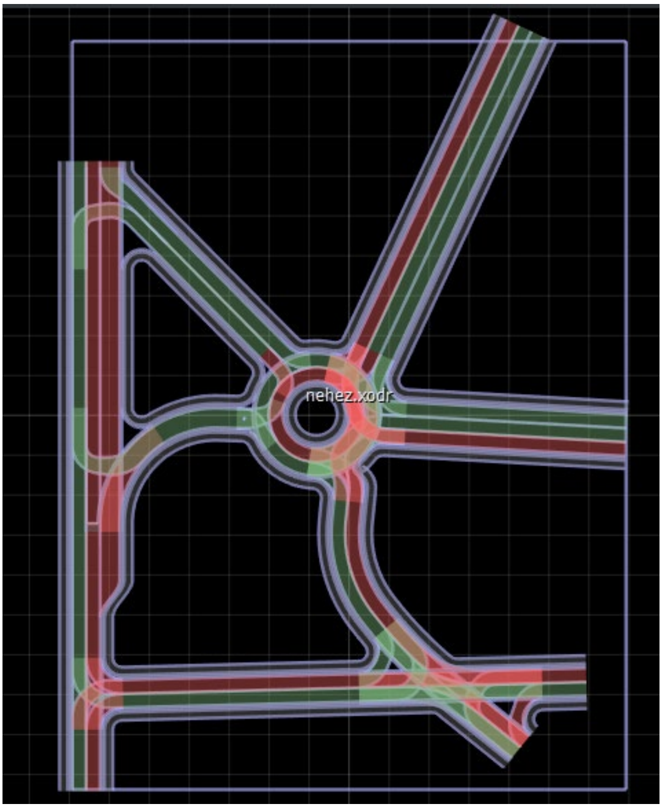


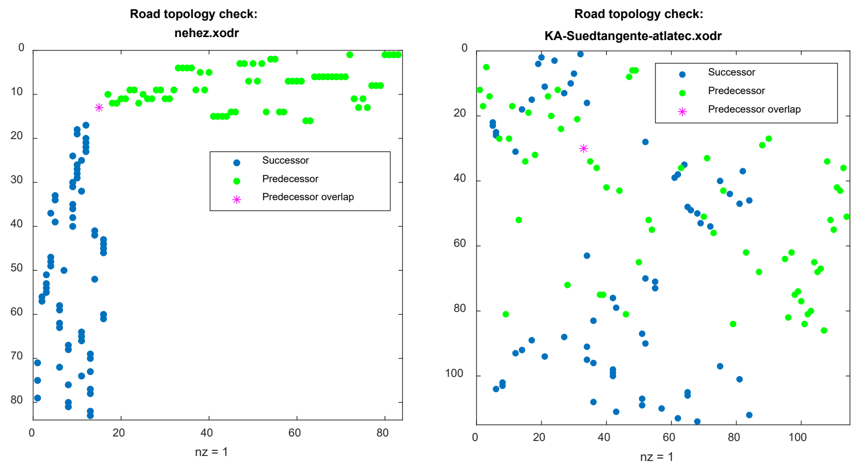
Figure 9
Synthetic OpenDRIVE sample for topological analysis

The sample area shown consists of 83 roads, 258 lanes, 7 intersections, and 117 lanes within them. To study the OpenDRIVE topology, 10 basic rules have been established [9]. An extracted example of the rules: using adjacency matrix equation (2) and incidence matrix equation (3), the following theorem should hold for the endpoints of the path reference lines:

$$\sum_j B_{ij}^{EP} = \sum_j A_{kj} \tag{7}$$

where B_{ij}^{EP} is the endpoint focused incidencies (edge-vertex relations).

OpenDRIVE distinguishes between predecessors and successors of roads and lanes, which are contained in links and junction elements (see Section 3). In addition, inference rules can be defined for these to check the consistency of the database [18]. An excellent example of inference rules [19] is that if for two consecutive roads it is true that the successor of Road 2 is Road 1 and the endpoint of Road 1 is the same as the starting point of Road 2 (a variant of a possible continuation), then the predecessor of Road 2 must be Road 1. Furthermore, the conclusions so applied can be extended to the numbered lanes of the roads in a sequence, allowing each lane to be uniquely identified and their junctions to be examined. In ambiguous cases, OpenDRIVE requires the use of the standard junction. However, it is very important from a topological point of view that the inference rules can be used to examine the junction lanes and compare them with the links. If they do not match, consistency errors occur, which are reported by the simulators as continuity errors. The synthetic example has one case of a predecessor overlapping error, which is currently only available as an error list and can be corrected manually (Fig. 10a).



a) synthetic example

b) Karlsruhe real test site

Figure 10

Predecessor-successor verification using synthetic and real examples

In parallel, we also carried out tests with real data. Thus, several models derived from field surveys (e.g., by Atlatec [20]), such as from German motorways, expressways or urban road networks, were analyzed. One of these surveyed models is the motorway near Karlsruhe ("Südtangente"), which consists of 114 roads, 886 lanes, of which 189 are junction lanes and 22 junctions. It is interesting to note that the XODR model from the German A9 motorway survey for international practice is completely flawless, although containing 248 roads, 2 683 lanes, 510 junction lanes, and 63 junctions (Fig. 10b).

Conclusions

In this work, we have reinforced the idea that topology is an extremely important part of the map content in geoinformatics, especially in the world of vehicles and transportation. We examined four different formats in which topology is implemented in different ways. These are, ESRI coverage, OpenStreetMap, Navigation Data Standard and OpenDRIVE respectively. In the course of the research, we found that topology can be embodied in quite different interpretations in these formats. Accordingly, we tailored our investigations to the data of each implementation: for OSM we focused on the large network connectivity, for NDS we looked at the lane geometry and connectivity appropriate for HD map content, and for OpenDRIVE we looked at the lane connectivity essential for smoothly serving simulations.

Our research required different tools: parsing XML files, populating and interpreting custom topological tables, graph analyzing techniques, establishing topological and inference rules and applying the rule set – to name just a few of the solutions implemented. We have shown that topology, which has been largely neglected, until now, is crucial and needs to be further investigated, verified and sometimes, fixed. Our overall aim is to create a procedure that is as automated as possible and performs the discussed steps, with minimal human intervention. We are therefore designing such a model application, which we intend to implement with the inclusion of Artificial Intelligence.

Acknowledgment

The research reported in this paper and carried out at the Budapest University of Technology and Economics has been supported by the National Research Development and Innovation Fund (TKP2020 Institution Excellence Subprogram, Grant No. BME-IE-MIFM) based on the charter of bolster issued by the National Research Development and Innovation Office under the auspices of the Ministry for Innovation and Technology. The project has been supported by the European Union, co-financed by the European Social Fund. EFOP-3.6.3-VEKOP-16-2017-00001. Supported by the ÚNKP-21-3 New National Excellence Program of the Ministry for Innovation and Technology from the source of the National Research, Development and Innovation Fund.

References

[1] M. Burg, “Simulation as Tool,” 2019

[2] “Mechanical Simulation.” [Online] Available: https://www.carsim.com/ [Accessed: 26-Sep-2019]

[3] “VTD - VIRES Virtual Test Drive.” [Online] Available: https://vires.com/vtd-vires-virtual-test-drive/ [Accessed: 26-Sep-2019]

[4] Vires, “Vires VTD,” 2022 [Online] Available: https://vires.mscsoftware.com/ [Accessed: 02-Apr-2022]

[5] “Simulation Technologies - avl.com.” [Online] Available: https://www.avl.com/hu/web/guest/simulation [Accessed: 26-Sep-2019]

[6] “SUMO - Simulation of Urban Mobility.” [Online] Available: http://sumo.sourceforge.net/ [Accessed: 26-Sep-2019]

[7] M. T. Horváth, Q. Lu, T. Tettamanti, Á. Török, and Z. Szalay, “Vehicle-In-The-Loop (VIL) and Scenario-In-The-Loop (SCIL) Automotive Simulation Concepts from the Perspectives of Traffic Simulation and Traffic Control,”

Transp. Telecommun. J., Vol. 20, No. 2, 2019

[8] "CarMaker | IPG Automotive." [Online] Available: https://ipg-automotive.com/products-services/simulation-software/carmaker/ [Accessed: 26-Sep-2019]

[9] J. M. Lógó and A. Barsi, "THE ROLE OF TOPOLOGY IN HIGH-DEFINITION MAPS FOR AUTONOMOUS DRIVING," Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci., Vol. XLIII-B4-2022, pp. 383-388, Jun. 2022

[10] ESRI, "Geodatabase topology rules and fixes for polyline features—ArcGIS Pro | Documentation." [Online] Available: https://pro.arcgis.com/en/pro-app/2.8/help/editing/geodatabase-topology-rules-for-polyline-features.htm [Accessed: 25-Apr-2022]

[11] OpenStreetMap, "OpenStreetMap," 2022 [Online] Available: https://www.openstreetmap.org. [Accessed: 25-Apr-2022]

[12] OpenStreetMap, "Stats – OpenStreetMap Wiki," 2022 [Online] Available: https://wiki.openstreetmap.org/wiki/Stats. [Accessed: 25-Apr-2022]

[13] OpenStreetMap, "Planet.osm – OpenStreetMap Wiki," 2022 [Online] Available: https://wiki.openstreetmap.org/wiki/Planet.osm [Accessed: 25-Apr-2022]

[14] NDS, "Navigation Data Standard - Wikipedia," 2022 [Online] Available: https://en.wikipedia.org/wiki/Navigation_Data_Standard [Accessed: 25-Apr-2022]

[15] NDS, "Navigation Data Standard (NDS) - The worldwide standard for map data in automotive eco-systems," 2022 [Online] Available: https://nds-association.org/. [Accessed: 25-Apr-2022]

[16] HERE, "HD Maps for Autonomous Driving and Driver Assistance | HERE," 2018 [Online] Available: https://www.here.com/products/automotive/hd-maps. [Accessed: 08-Dec-2019]

[17] MathWorks, "RoadRunner - MATLAB & Simulink." [Online] Available: https://www.mathworks.com/products/roadrunner.html [Accessed: 02-Apr-2022]

[18] M. Barsi and A. Barsi, "TOPOLOGICAL ANOMALY DETECTION IN AUTOMOTIVE SIMULATOR MAPS," in *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2022, p. 6

[19] Wikipedia, "Inference engine." [Online] Available: https://en.wikipedia.org/wiki/Inference_engine [Accessed: 02-Apr-2022]

[20] A. GmbH, "Atlatec," 2022 [Online] Available: https://atlatec.de/en/ [Accessed: 02-Apr-2022]