

AUTOMATIC EXTRACTION of IndoorGML CORE MODEL from OpenStreetMap

S. S. Mirvahabi ^a, R. A. Abbaspour ^b

^{a,b} School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran, Iran
{mirvahabi.s, abaspour}@ut.ac.ir

Commission VI, WG VI/4

KEY WORDS: IndoorGML, OpenStreetMap, Navigation, OGC, VGI

ABSTRACT

Navigation has become an essential component of human life and a necessary component in many fields. Because of the increasing size and complexity of buildings, a unified data model for navigation analysis and exchange of information. IndoorGML describes an appropriate data model and XML schema of indoor spatial information that focuses on modelling indoor spaces. Collecting spatial data by professional and commercial providers often need to spend high cost and time, which is the major reason that VGI emerged. One of the most popular VGI projects is OpenStreetMap (OSM). In this paper, a new approach is proposed for the automatic generation of IndoorGML data core file from OSM data file. The output of this approach is the file of core data model that can be used alongside the navigation data model for navigation application of indoor space.

1. INTRODUCTION

In the recent years, navigation has become an essential component of human life and a necessary component in many fields such as disaster management, traffic management, and urban planning. Because of the increasing size and complexity of buildings, navigation is not only restricted to the outdoor environment but also covers indoor environment (Kim et al., 2013). A large number of applications, which are based on routing inside public buildings such as hospitals, shopping malls, universities, and airports, have been created. This type of applications needs the same data model for navigation analysis and exchange of information to each other (Li et al., 2010). Converting spatial data to a unified format is a discussable topic in navigation. IndoorGML data model is the proposed data structure for this issue. On the other hand, indoor navigation needs comprehensive geographic data for analyses and applications. Projects such as OSM often provide the cheapest source of geospatial data that is useful for outdoor and indoor navigation.

In the past years, several research was focused on indoor navigation problems based on VGI data. Goetz discussed the automatic generation simulation-related data based on OSM (IndoorOSM) to perform multi-agent evacuation simulations (Goetz et al., 2012a). Mortari presented the automatic extraction of the connectivity graph of a building from CityGML structure for indoor navigation (2013). Herrera et al. (2014) used IndoorOSM data and hardware component of smartphone to provide a hybrid indoor localization approach. Zaragozí et al. proposed a hybrid navigation system based on open data and published via standard protocols and web services (Zaragozí et al., 2015). However, to the authors of this paper's knowledge there is currently no an approach for automated extraction IndoorGML data from OSM data.

The main purpose of this paper is development of an automatic approach for automatic generation of IndoorGML data core file from OSM data file. The rest of this paper is organized as follows. In Section 2, OSM and IndoorOSM are introduced. In Section 3, the IndoorGML standard is described in details required for the discussion. This is followed by representing methodology and its implementation in Section 4. Conclusion is final section.

2. OPENSTREETMAP AND INDOOROSM

Navigation requires extensive, high-quality geospatial data for applications and analysis. Collecting spatial data by professional and commercial providers often need to spend high cost and time, which is the major reason that Volunteered Geographic Information (VGI) has become more attractive to professionals in the geo-information industry in the last few years. People participation without the need to specialized knowledge, self-improvement, self-promotion and creation dynamic, and cost effective spatial information infrastructure are benefits of the VGI (Goodchild, 2007).

One of the most popular VGI projects is OSM. The main goal of OSM is generation of a free online map that covers whole globe. There exist three levels of editors (iD, potlatch, and JOSM) that allow any user with any level of expertise to create, edit, and improve spatial data in OSM. Geometric data collected by GPS receivers or individual knowledge of people is shared in the form of node, way, and relation in OSM databases. Moreover, semantic information in the form of key-value pairs can be linked to geometric data. This project has been extended in many different directions, such as indoor mapping, 3D mapping, geocoding, and routing services.

At the end of 2011, OSM launched IndoorOSM, which is a tagging schema for indoor mapping following OSM

methodologies. It provides two dimensional geometries with additional semantic or metric indoor information (Goetz et al., 2012b). IndoorOSM represents floors, rooms, corridors, doors, and POIs of indoor environments by using nodes, ways, and relations. All map data from OSM can be exported as an XML *.osm format for further processing.

3. IndoorGML STANDARD

CityGML, KML (Keyhole Markup Language), and IFC (Industry Foundation Classes) are several standards that focus on 3D geometry and semantic of buildings but they lack necessary components that are needed by indoor navigation applications. IndoorGML standard describes an appropriate data model and XML schema of indoor spatial information that focuses on modelling indoor spaces for navigation goals. It is implemented as an application schema of the Geography Markup Language (GML) version 3.2.1. This standard comprises of two basic components: a core data model (to describe topological connectivity and different contexts of indoor spaces) and a data model for navigation in indoor spaces (OGC, 2012). Indoor of buildings can be modeled as cellular spaces by non-overlapping cells. Each cell is defined as the smallest unit of a building such as rooms and corridors (Mortari, 2015) and has unique ID and common boundary. A cellular space S is defined as follows:

$$S = \{c_1, c_2, c_3, \dots, c_n\} \quad (1)$$

where c_i is the i^{th} cell (OGC,2012).

In IndoorGML structure, it is required to describe the topological relationships expressly. The Node Relation Graph (NRG) is generally used for building data to

represent topological relationships, e.g. adjacency and connectivity among buildings (Lee, 2004). Furthermore, the Poincaré duality (Munkres, 1984) is used for representing topological relationships. Figure 1 indicates the graphic representation of NRG graph in the building plan for part of the case study.

Because the geometric representation of features in indoor space is not a major focus of IndoorGML, there are three ways for representing geometry of building parts in IndoorGML data model (OGC, 2012):

1. External reference: An IndoorGML document only contains external links to objects defined in other data sets such as CityGML file.
2. Geometry in IndoorGML: Geometric representation of each building part may be included within an IndoorGML document. It is GM_Surface in 2D and GM_Solid in 3D spaces as defined in ISO 19107 (ISO 19107, 2003).
3. No geometry: No geometric information is also included in IndoorGML document.

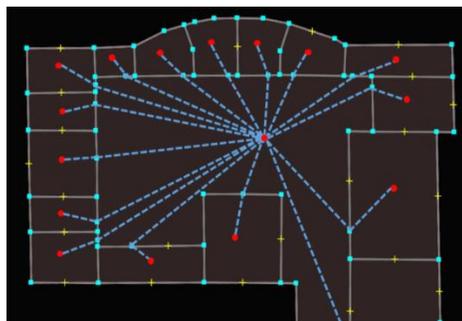


Figure 1. NRG graph

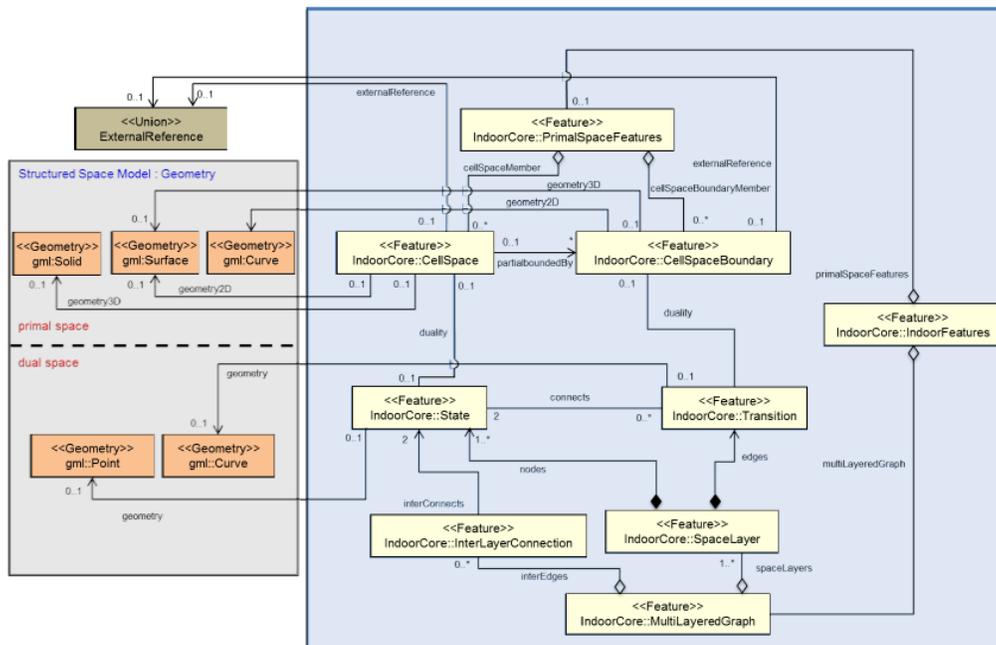


Figure 2. IndoorGML core module data model (OGC, 2012)

The UML diagram depicted in Figure 2 shows data model of core module for IndoorGML. The model consists of two main classes: *PrimalSpaceFeatures* and *MultiLayeredGraph*. The Classes *CellSpace* and *CellSpaceBoundary* are two subclass of class *PrimalSpaceFeatures* that demonstrate the geometry and semantic of cellular space according to the concept of geographic features (ISO 19109). To define one cell in Euclidean primal space, *CellSpace* class is represented as *gml:Solid* or *gml:Surface* in 3D and 2D space, respectively. *CellSpaceBoundary* is also used as *gml:Surface* or *gml:Curve* to semantically describe 2D/3D boundary of each cell.

The class *MultiLayeredGraph* consists of several *SpaceLayer*. The multi-layered space model allows for the coherent combination of different decompositions of space according to different semantics. Each *SpaceLayer* demonstrates the topological relation between cells (building parts) that is derived from the NRG graph. *States* and *Transitions*, are also used as dual representations of cells in primal space.

4. METHODOLOGY AND IMPLEMENTATION

Our case study in this paper is the building for school of surveying and Geospatial Information Engineering at University of Tehran. The plan of case study building, which already exist in OSM, is shown in Fig. 4.

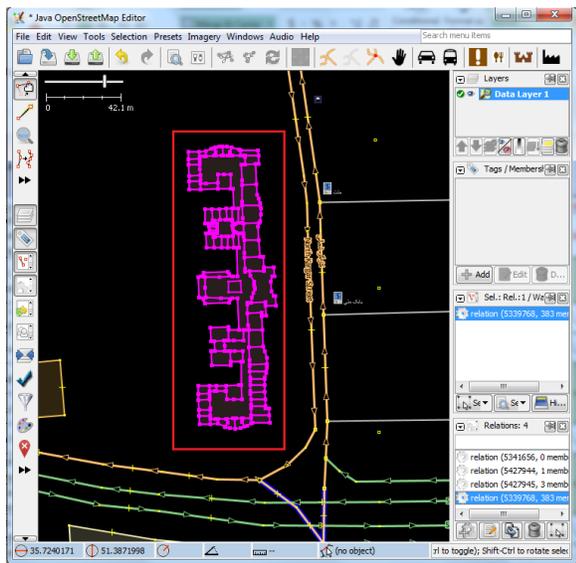


Figure 3. Footprint of case study building in JOSM

The proposed methodology is illustrated in Figure 5. To implement this approach, firstly, the building footprint for university of Tehran with floors was entered as relation in the OSM via JOSM editor (IndoorOSM). Besides, for each building part, semantic information such as part ID, room, name, and building level was added. Locations of door and entrance were also added. Then, footprint of mentioned building was exported to *.osm format from OSM portal and was used as input data for the approach. The exported *.osm file contains a data model that consists of node, way,

relation, and semantic tag data of building. There is a hierarchical relationship between members of file and each member that has a unique ID and its semantic tag. This structure was read by proposed approach; required elements of it was extracted and saved in the suitable data model.

NRG is used for storing topology of building in IndoorGML. Our approach calculated the centroid (Eq.2 and Eq.3) of each building part and coordinates of each door, saved them as nodes of duality structure and then created the NRG graph from these nodes.

$$A = \frac{1}{2} \sum_{i=0}^{N-1} (x_i y_{i+1} - x_{i+1} y_i) \quad (2)$$

$$C_x = \frac{1}{6A} \sum_{i=0}^{N-1} (x_i + x_{i+1})(x_i y_{i+1} - x_{i+1} y_i) \quad (3)$$

$$C_y = \frac{1}{6A} \sum_{i=0}^{N-1} (y_i + y_{i+1})(x_i y_{i+1} - x_{i+1} y_i)$$

where A is the area of polygon, N is the number of polygons point, x_i and y_i are coordinates of i^{th} vertex of polygon, and C_x and C_y are coordinates of centroid of polygon.

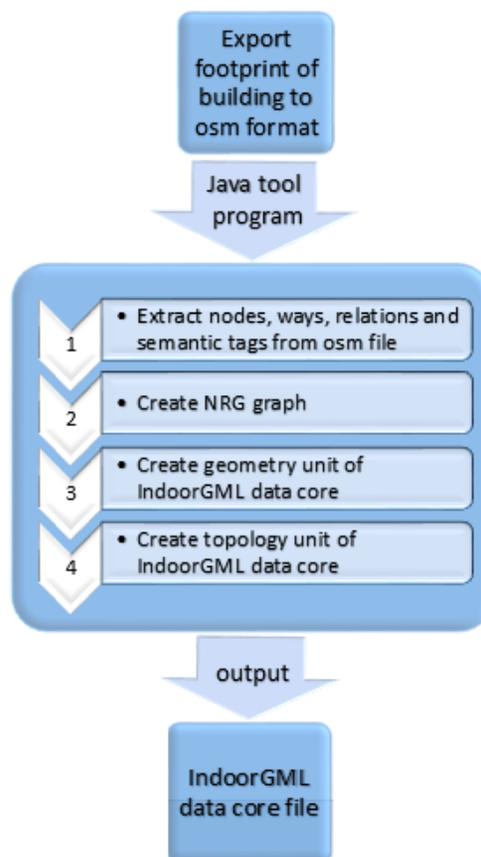


Figure 4. Proposed methodology

To create PrimalSpaceFeatures unit of IndoorGML core model, we used 2D geometry derived from *.osm nodes, ways, relations, and tags data. Each building part is defined as *gml:Surface* in class *CellSpace* with *gml:Curve* boundary in class *CellSpaceBoundary*.

In MultiLayeredGraph unit, we have one *SpaceLayer*, which shows the topological relation of navigation layer based on NRG graph. *SpaceLayer* consists of two subunits: *nodes* (set of *States*) and *edges* (set of *Transition*). The connected *transition* and dual part is defined for each *State*. The connected *State* are also defined for each *Transition*. Figure 4 indicates parts of the output file, which shows one *State*.

```
-<stateMember>
- <State gml:id="R3">
  <gml:name>001</gml:name>
  <duality xlink:href="#C3"/>
  <connects xlink:href="#T1"/>
  <connects xlink:href="#T2"/>
  <connects xlink:href="#T3"/>
  <connects xlink:href="#T4"/>
  <connects xlink:href="#T5"/>
  <connects xlink:href="#T6"/>
  <connects xlink:href="#T7"/>
  <connects xlink:href="#T8"/>
  <connects xlink:href="#T9"/>
  <connects xlink:href="#T10"/>
  <connects xlink:href="#T11"/>
  <geometry>
  - <gml:Point gml:id="P3">
    <gml:pos>445538.543473167 5444902.27372664 -2.02</gml:pos>
    <gml:Point>
    </gml:Point>
  </geometry>
  </State>
</stateMember>
```

Figure 5 Part of output file

The output of our approach is a core data model file of IndoorGML. This approach is written in Java language and can convert OSM data file automatically to XML structure of IndoorGML data core model.

5. CONCLUSION AND FUTURE WORK

In this paper, we proposed a java-based approach for automatic extraction of IndoorGML and the VGI data of building for school of surveying and geospatial information engineering at university of Tehran is used as the case study. In the first stage, it extracted geometric and semantic data from OSM file and then calculated the NRG graph from these data. Then our approach provided geometrical and topological unit of IndoorGML data core. The output of this approach is a core data model file that can be used alongside the navigation data model for navigation application of indoor space. This output file is in XML format and has two main units: geometry unit and topology unit. We created the core data model for one building in our case study but this approach can be developed to create data core model from several building and combined with outdoor data model.

References

- Goetz, M. and Zipf, A, 2012a. Towards defining a framework for the automatic derivation of 3D CityGML models from volunteered geographic information. *Int. J. 3-D Inf. Model. 1*, 1–16.
- Goetz, M., and Zipf, A, 2012b. Mapping the Indoor World: Towards Crowdsourcing Geographic Information About Indoor Spaces. *GIM international* 26.3, pp. 30–34.
- Goodchild, M.F, 2007. Citizens as sensors: the world of volunteered geography. *GeoJournal* 69 (4), pp. 211–221.
- Herrera, J. A, Plöger, P., Hinkenjann, A. Maiero, J., Flores, M. and Ramos, A., 2014. Pedestrian indoor positioning using smartphone multi-sensing, radio beacons, user positions probability map and IndoorOSM floor plan representation. In *Proceedings of the International Conference on Indoor Positioning and Indoor Navigation (IPIN '14)*, pp. 1–10, Busan, Republic of Korea.
- ISO 19107, 2003. Geographic information – Spatial Schema.
- Kim, Y.-J., Kang, H.-Y and Lee, J., 2013. Development of indoor spatial data model using CityGML ADE. In *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pp. 41–45.
- Li, K. J., Lee, J., 2010. Indoor spatial awareness initiative and standard for indoor spatial data. *Proc IROS*, pp.76–80.
- Lee, J., 2004. A Spatial Access Oriented Implementation of a Topological Data Model for 3D Urban Entities, *Geoinformatica* 8(3), pp. 235–262.
- Mortari, F, 2013. Automatic Extraction of Improved Geometrical Network Model from CityGML for Indoor Navigation, *Master Thesis*.
- Munkres, J. R., 1984. *Elements of Algebraic Topology*, Addison-Wesley, Menlo Park, CA.
- OGC, 2012. OGC® IndoorGML Encoding Standard, Version: 1.0, OGC Doc. No. 08-007r1.
- Zaragozí, B., Giménez-Font, P., Navarro-Carrion, J., T., Rico-Amorós, A., Ramón –Morte, A. and Belda-Antoi, A., 2015. A Proposal for Design and Implementation of an Hybrid Navigation System Based on Open Data, augmented Reality and Big Data Applications for the Smart Cities. In *Proceedings of the 1st International Conference on Geographical Information Systems Theory, Applications and Management*, 221–226, Barcelona, Spain.