

TOWARDS GENERATING SEMANTICALLY-RICH INDOORGML DATA FROM ARCHITECTURAL PLANS

Srishti Srivastava^a, Nishith Maheshwari^b, K. S. Rajan^c

Lab for Spatial Informatics, IIIT Hyderabad, India -
(^aar.srishti@gmail.com, ^bnishith.maheshwari@research.iiit.ac.in, ^crajan@iiit.ac.in)

Commission IV, WG IV/5

KEY WORDS: CAD drawing, data conversion, IndoorGML, spatial modeling, semantics, topology

ABSTRACT:

Recent years has seen an increase in the work done on indoor data mapping and modeling. The standard data models provide different ways to store and access the indoor data but the way it is done is specific to the domain in which they are used. Although models like IFC, CityGML and IndoorGML provides rich functionality, the widespread availability of indoor data is not in these formats. This paper presents a step by step methodology to convert indoor building data of existing buildings, represented in architectural drawings into a topologically consistent and semantically rich indoor spatial model. The workflow presented consists of extracting relevant geometric entities from CAD drawings, assessing their topological relationships, using it to derive semantic information of spaces and making the data available in the form of IndoorGML. Since the current IndoorGML features lack the capability to store relevant semantic information, a semantic extension to IndoorGML is also proposed. The extraction of primitive spatial elements in rectilinear buildings like walls and doors are considered for the work presented in this paper. Development of a toolkit which implements this methodology in a seamless manner is work in progress and would incorporate extraction of complex spatial elements like staircases, ramps, curvilinear walls and windows, which is out of scope of the current work presented in this paper.

1. INTRODUCTION

Today's commercial CAD software is the product of years of research that began in the 1960's and 1970's. These applications have found widespread use in the architectural marketplace; nevertheless they represent only the first fruits of research in computer graphics largely focussed on representation of architectural elements to aid AEC. The huge advancement of IT including the enhanced processing power and availability of hand-held devices which produce huge amount of spatial data in the last decade calls for cutting edge research in indoor mapping domain.

One of the primary purpose of CAD architectural plans is for designing and drafting the elements of a building using a representational standard symbology. In general these representational standards are visual in nature and are meant to be interpreted manually by professionals (Cherneff et al., 1992). CAD drawings capture a rich amount of indoor information of a building using a representational standard which is in the form of primitive geometric elements like points, lines and arcs. These elements are stored in an incoherent manner in a CAD drawing which results in missing well defined indoor geometric spaces and topology. Also there is no global coordinate reference system which creates a major hurdle for CAD drawings to be translated to a GIS based indoor model. Thus a need for digital interpretation process to effectively capture the components of the building is necessary.

Various research works have been done to extract meaningful information from architectural plans. Some works have focused on 3D reconstruction of building from 2D CAD drawings (So et al., 1998, Lu et al., 2005, Yang et al., 2008) with the purpose of visualisation while others have focused on extracting building elements from scanned floorplans to create 3D models and store as BIM IFC format (Gimenez et al., 2016) or as IndoorOSM (Wu,

2015). Work done by (Domínguez et al., 2012) has suggested a semi-automatic solution for detecting floor topology in the context of walls only for the purpose of extrusion. Although they are all able to capture individual basic building entities and create a 3D solid model, the topological relationships between the spaces are missing.

(Huang et al., 2008) and (Lewis and Séquin, 1998) have defined a workflow for recognising enclosed spaces from architectural CAD drawings and defining their topological relationships with openings. They have also talked about assigning semantic information based on function of the spaces. But the output from these works is not available in a standard data format which can be used for various application scenarios. Hence a need for the data to be available in a standard which incorporates the geometric spaces, their topological relationships and semantic information is there.

Current prevalent standards for describing indoor spaces are CAD drawings, BIM IFC model, CityGML Indoor ADE (Application Domain Extension) and IndoorGML -

- Drawings have been used by architects and designers to define the layout of the building along with functional use of the spaces. It is one of the most widely used and available standard to define a buildings specification
- IFC (Industry Foundation Classes), an open standard for BIM (Building Information Model), is developed and supported by BuildingSMART organisation and its primary purpose is to provide interoperability between the AEC (Architecture, Engineering and Construction) industry (buildingSMART International Ltd., 2018). It is used for estimating the construction material for the buildings, facility management etc.

- CityGML with the Indoor ADE is used to describe the indoors in different level of details (LODs) (Gröger et al., 2012). LOD3 describes the indoors as boundary surfaces. LOD4 can be used to describe the furniture and other indoor elements
- IndoorGML is an OGC standard for defining indoor spatial information with specific focus on indoor navigation (Li et al., 2015)

	CAD Drawing	BIM IFC Model	CityGML	IndoorGML
Primary Purpose	To represent 2D design data and convey the design elements to the AEC team during the construction process	To describe building and construction industry data and facilitate interoperability between architecture, engineering and construction industry. Used mainly for calculating the amount of material to be used for construction	To reach a common definition of the basic entities, attributes, and relations of a 3D city model. The model not only focus on visualisation but it is possible to do thematic queries, analysis of tasks and mining of spatial data	To represent and allow for exchange of geo-information required to build and operate indoor navigation systems
Representation of Navigational network	No	No	No	Yes. There is a navigation module in IndoorGML
Relationships stored	No	Has relationships like assigns, associates, connects and defines between the objects	Relationships in the form of aggregations, generalisations and associations are stored	Topological properties like connectivity, adjacency, accessibility in the form of Node-Relation Graph (NRG)

IFC and CityGML provides a good model for representing the building entities in a queryable model. But the availability of real world data in these models is not very prevalent. IndoorGML model provides the most functionality and flexibility among these standards. It provides a flexible way to store the geometry either in the model itself by using GML notation or by referencing external data sources (like IFC or CityGML). It also provides extensive support for indoor navigation modeling via its Navigation module.

This paper presents a workflow which focuses on extracting the relevant geometric entities from CAD drawings, assessing their topological relationships, using it to derive semantic information of spaces and making the data available in the form of IndoorGML. The current IndoorGML features lack the capability to store relevant semantic information, hence a semantic extension to IndoorGML is also proposed. The primitive elements considered for extraction is limited to only walls and doors. For the work presented in this paper only rectilinear buildings are considered and walls represented by polylines or curved lines are out of the scope of this paper work.

In Section 2, details about each step of our methodology and the proposed semantic extension to IndoorGML is explained. Finally in the last section we talk about the implication of this work and discussions for the the future.

2. METHODOLOGY

The objective of our work is to use the information available in architectural drawings (2D CAD drawings) to generate an indoor model containing the geometric, topological and semantic information of the building. The input for our workflow is architectural presentation drawing in the form of CAD DXF file representing one floor of a building.

Our workflow comprises of the steps shown in Figure 1 and can be summarized as follows -

Step 1: Primitive element extraction and interpretation.

Step 2: Generation of space geometries and creation of topological relationships.

Step 3: Semantic classification of spaces based on geometry and topology.

Step 4: Finally the output in the form of IndoorGML files. An extension to IndoorGML is proposed inorder to store the semantic information.

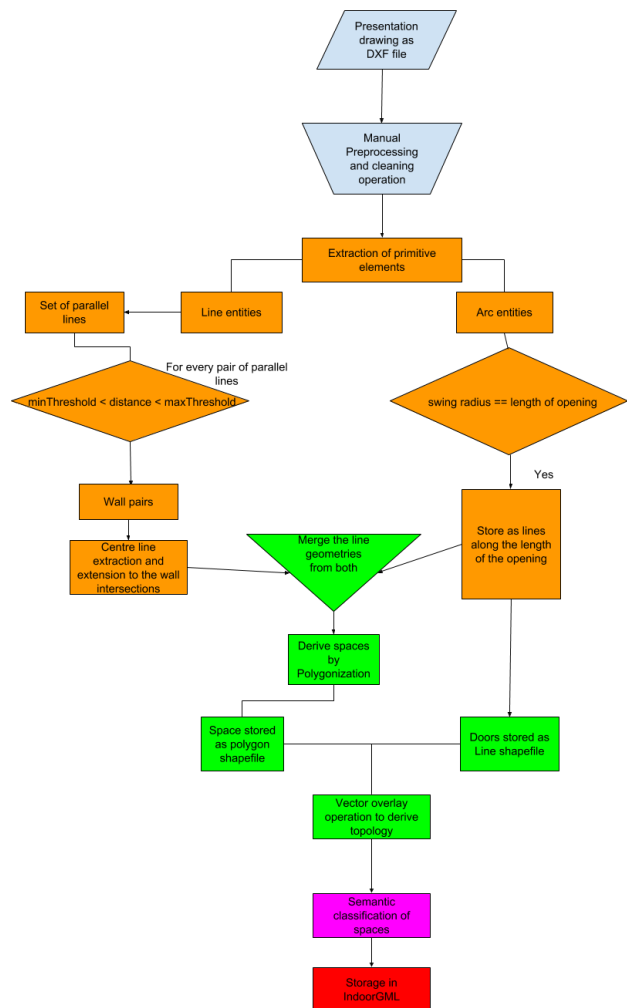


Figure 1. Flowchart of the methodology. Each color represents different phase of the workflow. (a)Orange: Element extraction and interpretation, (b)Green: Space geometry and topology generation, (c)Pink: Semantic classification, (d)Red: Storage of data as semantically-rich IndoorGML

2.1 Element extraction and interpretation

Any drawing is composed of different spatial elements which come together to specify the building created by the architect/designer. Interior spaces are represented in architectural drawings

using the primitive spatial elements which are - walls, doors, windows, corridors and staircases. These elements are represented in a drawing using a standard symbology.

Spatial elements are of two major types:

- *Primitive Spatial Elements* : There are some primitive spatial elements like walls, windows, doors which are either opaque or transparent. These elements are used in defining a space and its extents.
- *Complex Spatial Elements* : These spatial elements are responsible to connect either one space to another or an interior space to an exterior space, at one level or different levels. Complex Spatial Elements are used to connect spaces, defined using primitive elements. The connector elements like corridors are for horizontal connectivity between spaces on the same plane or vertical connectors like Staircase, Ramps, Elevators, Escalators across different planes from one floor to other floors.

In the methodology, the extraction of the primitive elements is restricted to only walls, doors and text labels present in the DXF file. Since we have considered only rectilinear spaces and basic definition of a door, only the walls represented by parallel lines and doors represented by arc are extracted. Cases with curvilinear walls and complex spatial elements are not handled currently.

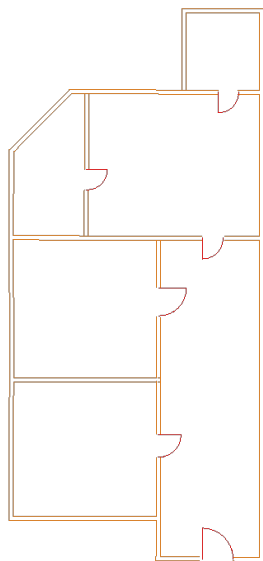


Figure 2. Sample CAD floorplan input

2.1.1 Wall extraction : A wall is one of the basic component which makes up the space in a building. For our work we have only considered the walls represented by parallel lines. The summary of the steps involved is as follows -

- Extraction of all line entities from the DXF file
- Identification of wall pairs : This is done by identifying pairs of parallel lines which lie within a threshold distance of each other (as shown in Figure 3).
- Generation of center line corresponding to each wall pair
- Identifying the intersection of the walls by checking for every wall pair if they are intersecting other wall pairs (as shown in Figure 4).
- If intersection between two wall pairs is found then their center lines are extended to the point of intersection

The extraction of centre line of wall pairs is done to derive the spaces and maintain their topological consistency.

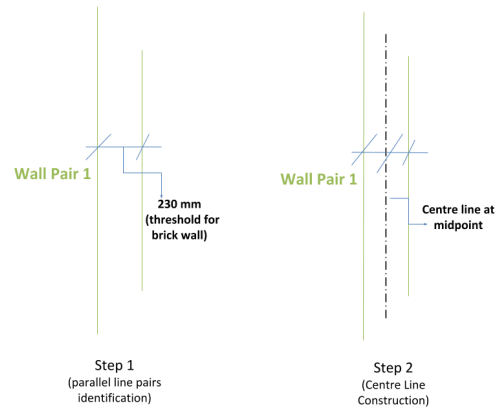


Figure 3. Identification of wall pairs and centreline generation

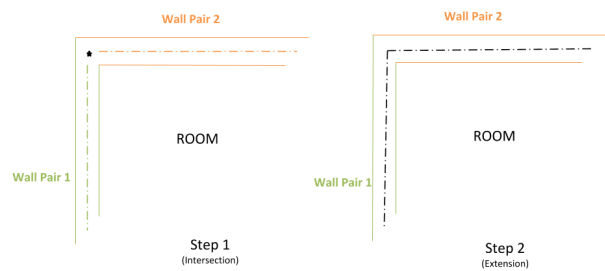


Figure 4. Centre line extension

2.1.2 Door extraction : A door/opening is an indoor element which connects different spaces to each other and hence play an important role when building the topology. The standard symbology as shown in the Figure 5 is used to interpret the door. It is identified by extracting the arc entities in the DXF file and retaining the ones which have the swing radius of the arc equal to the length of the opening.

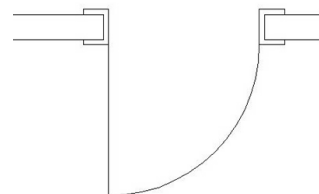


Figure 5. Symbology used to represent doors in a drawing

2.1.3 Text information : A lot of labels are associated with different elements in a drawing. These labels contain important semantic information about these elements like the function, name, level and other relevant properties. These are identified by extracting the text entities from the DXF file and associating them with the corresponding elements from the previous step.

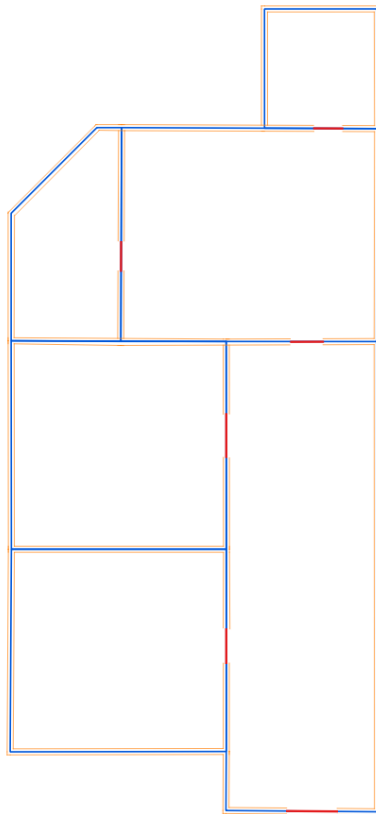


Figure 6. Output after wall and door extraction. The blue lines represent the wall extracted after centre line generation and red line segments represent the doors.

2.2 Space geometry and topology generation

Using the extracted walls and doors from the previous step, the enclosed spaces are derived by joining the threshold gaps and detecting loops. The spaces are then stored as polygons in the form of a shapefile. Each polygon represents a quantified space (usually a room, corridor) as defined in the original drawing. Similarly the doors are stored in the form of a line shapefile layer. The text information extracted from the labels are stored in the form of attributes to the corresponding features.

In order to perform meaningful spatial operations on the geometry, the topological relationships between the features need to be defined. For our work two kinds of relationships are analysed -

- *Adjacency or Borders* : Doors bordering a space. Since any numbers of doors can be present in a space, the space-door mapping is in the ratio 1:n
- *Connectivity* : Spaces connected by a door. A door can connect two spaces to one another, so the door-space mapping is in the ratio 1:2

These relationships are derived by applying vector overlay operations on the space and door shapefiles.

2.3 Semantic classification

With the help of the geometry and the topological relationship between the elements, the semantic information like the class and

capacity is derived (Maheshwari and Rajan, 2016). The capacity of a space is calculated using area from the geometric information and the class of the space is derived using the following rules -

- Spaces having 2 or less doors : Room
- Spaces with more than 2 doors and having L/B ratio (length by breadth) ≈ 2 : Corridor
- Spaces with more than 2 doors : Junction

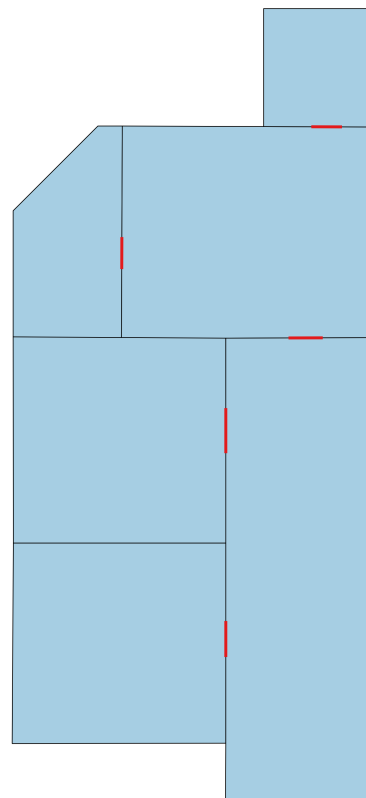


Figure 7. Output after polygon and topology generation

2.4 Semantically rich IndoorGML

For the purpose of our work we have considered only the Thin door model of IndoorGML since the doors are represented as lines. The indoor spaces are stored as CellSpace features and the doors are stored as CellSpaceBoundary features along with their geometry.

Each CellSpace feature and CellSpaceBoundary feature corresponds to a State feature and Transition feature respectively. The State feature has a point geometry (usually the centroid of the space) and stores the topological information as a list of Transition features it connects to. Similarly Transition feature stores the information of the two states it connects and the geometry is in the form of an edge with the states joined by the centre point of the opening.

The current structure of IndoorGML is insufficient to incorporate the semantic information of the entities. The type of semantic

information stored is decided on the basis of the semantic model proposed in (Maheshwari and Rajan, 2016). Hence an extension to the IndoorGML Core module is proposed in order to include the semantic information in IndoorGML.

Following is a snippet of the proposed semantic extension -

```
<xs:element name="SemanticSpace" type="SemanticSpaceType"
  substitutionGroup="IndoorCore:CellSpace">
  <xs:annotation>
    <xs:documentation>SemanticSpace
  </xs:documentation>
  </xs:annotation>
</xs:element>

<xs:complexType name="SemanticSpaceType">
  <xs:complexContent>
    <xs:extension base="IndoorCore:CellSpaceType">
      <xs:sequence>
        <xs:element name="name" type="xs:string"/>
        <xs:element name="class" type="gml:CodeType"/>
        <xs:element name="capacity" type="xs:double" minOccurs="0" maxOccurs="1"/>
        <xs:element name="impedance" type="xs:double" minOccurs="0" maxOccurs="1"/>
      </xs:sequence>
    </xs:extension>
  </xs:complexContent>
</xs:complexType>
<!-- ----->
<xs:element name="SemanticOpening" type="SemanticOpeningType"
  substitutionGroup="IndoorCore:CellSpaceBoundary">
  <xs:annotation>
    <xs:documentation>SemanticOpening
  </xs:documentation>
  </xs:annotation>
</xs:element>

<xs:complexType name="SemanticOpeningType">
  <xs:complexContent>
    <xs:extension base="IndoorCore:CellSpaceBoundaryType">
      <xs:sequence>
        <xs:element name="name" type="xs:string"/>
        <xs:element name="class" type="gml:CodeType"/>
        <xs:element name="capacity" type="xs:double" minOccurs="0" maxOccurs="1"/>
        <xs:element name="state" type="xs:string"/>
        <xs:element name="access" type="xs:string"/>
      </xs:sequence>
    </xs:extension>
  </xs:complexContent>
</xs:complexType>
```

Where class would denote the type of space or boundary (door, window, room, corridor), capacity would give the number of people the space can hold or the boundary can allow to pass at a time and impedance is calculated based on the obstacles in the space. The state element denotes whether the opening is open, closed or blocked and the access element defines what kind of access is there for the opening (lock, rfid or biometric authentication).

Following snippet shows how the data is stored in a semantically-rich IndoorGML-

```
<SemanticSpace gml:id="C1">
  <duality xlink:href="#S1"/>
  <cellSpaceGeometry>
    <Geometry2D>
      <gml:Polygon>
        <gml:exterior>
          <gml:LinearRing>
            <gml:pos>55.0869297445 687.594882353 0.0</gml:pos>
            <gml:pos>112.054835236 744.562787844 0.0</gml:pos>
            <gml:pos>128.450476329 744.562787844 0.0</gml:pos>
            <gml:pos>127.894691885 601.726185782 0.0</gml:pos>
            <gml:pos>55.0869297445 602.83775467 0.0</gml:pos>
            <gml:pos>55.0869297445 687.594882353 0.0</gml:pos>
          </gml:LinearRing>
        </gml:exterior>
      </gml:Polygon>
    </Geometry2D>
  </cellSpaceGeometry>
  <name value="ServerRoom"/>
  <class value="Room"/>
  <capacity>10</capacity>
</SemanticSpace>
<!-- ----->
<SemanticOpening gml:id="CB1">
  <duality xlink:href="#T1"/>
  <cellSpaceBoundaryGeometry>
    <Geometry2D>
      <gml:LineString>
        <gml:pos>256.458360412 744.066025222 0.0</gml:pos>
        <gml:pos>275.083297891 743.993747074 0.0</gml:pos>
      </gml:LineString>
    </Geometry2D>
  </cellSpaceBoundaryGeometry>
  <name value="MainEntrance"/>
  <class value="Door"/>
  <capacity>2</capacity>
  <state value="closed"/>
  <access value="rfid"/>
</SemanticOpening>
```

3. CONCLUSION AND FUTURE WORK

This paper presented a step by step methodology to convert indoor building data represented in architectural drawings into a topologically consistent and semantically rich indoor spatial model. This is done by extracting building elements like walls and doors from the architectural drawing, using them to generate topologically connected spaces, deriving semantic labels for spaces and finally storing the data in an extended IndoorGML model.

Development of a toolkit which implements this methodology in a seamless manner is under progress which would incorporate extraction of complex spatial elements (like staircases, ramps, curvilinear walls, windows), multiple floor handling and generation of navigation network. The usability of the output in application scenarios like spatial analysis, evacuation simulation and facility management will also be explored.

REFERENCES

buildingSMART International Ltd., 2018. Ifc overview summary. <http://www.buildingsmart-tech.org/specifications/ifc-overview>.

Cherneck, J., Logcher, R., Connor, J. and Patrikalakis, N., 1992. Knowledge-based interpretation of architectural drawings. *Research in Engineering Design*.

Domínguez, B., García, Á. L. and Feito, F. R., 2012. Semiautomatic detection of floor topology from CAD architectural drawings. *CAD Computer Aided Design*.

Gimenez, L., Robert, S., Suard, F. and Zreik, K., 2016. Automatic reconstruction of 3D building models from scanned 2D floor plans. *Automation in Construction*.

Gröger, G., Kolbe, T., Nagel, C. and Häfele, K., 2012. OGC city geography markup language (CityGML) encoding standard, version 2.0, ogc doc no. 12-019. *Open Geospatial Consortium*.

Huang, H. C., Lo, S. M., Zhi, G. S. and Yuen, R. K., 2008. Graph theory-based approach for automatic recognition of CAD data. *Engineering Applications of Artificial Intelligence*.

Lewis, R. and Séquin, C., 1998. Generation of 3D building models from 2D architectural plans. *CAD Computer Aided Design* 30(10), pp. 765–779.

Li, K.-J., Lee, J., Zlatanova, S., Kolbe, T. H., Nagel, C. and Becker, T., 2015. OGC® IndoorGML. *Open Geospatial Consortium* (05), pp. 1–17.

Lu, T., Tai, C. L., Bao, L., Su, F. and Cai, S., 2005. 3D reconstruction of detailed buildings from architectural drawings. *Computer-Aided Design and Applications* 2(1-4), pp. 527–536.

Maheshwari, N. and Rajan, K. S., 2016. A semantic model to define indoor space in context of emergency evacuation. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives* 41(July), pp. 313–318.

So, C., Baciú, G. and Sun, H., 1998. Reconstruction of 3D virtual buildings from 2D architectural floor plans. *Proceedings of the ACM symposium on Virtual reality software and technology*.

Wu, H., 2015. Integration of 2D architectural floor plans into Indoor OpenStreetMap for reconstructing 3D building models. Master's thesis, TU Delft, Delft University of Technology.

Yang, R., Lu, T. and Cai, S., 2008. 3D building reconstruction based on interpretation of architectural drawings. In: *Proceedings of the 2008 IEEE International Conference on Information and Automation, ICIA 2008*, pp. 1474–1479.