Topology-based Classification Error Calculation based on IndoorGML Document

Krisztián Ilku, Judit Tamás

Abstract: Topology-based classification error calculation method for symbolic indoor positioning is presented based on IndoorGML document. Symbolic indoor positions or Zones are well-defined parts of the building, which can be treated as a classification category. The evaluation of well-known classifiers is based on the classical CRISP approach, which considers each misclassification equally wrong. Our previous experimental results revealed the need to consider the topology in the classification error calculation. A possible solution for this challenge is gravitational force based approach, which calculates the classification error by the size and the layout of the Zones. Testing the criteria against this approach in real-life scenario, real-life environment is required. IndoorGML is a standard for specifying indoor spatial information, and it represents the indoor space as non-overlapping closed objects. These indoor spaces are bounded by physical or fictional boundaries, and the representation of an object is by both geometric shape and bounding box. Thus, IndoorGML standard can be used both for modeling the indoor environment and calculation the classification error for symbolic indoor positioning services. In this paper, the gravitational force based approach is examined in real-life environment of Institute of Information Science building in University of Miskolc defined in IndoorGML Document.

Keywords: IndoorGML, Classification Error, Indoor Localization, Topology

Introduction

Indoor Positioning System is considered as an active research field since the early 1990s, and these systems are detailed in the following surveys [1][2]. Absolute, proximity and symbolic indoor positioning are usually distinguished. Absolute location is the coordinates of an object. Proximity position is an orientation to known reference points. Symbolic position refers to the well-defined part of a building, such as rooms, and these positions can be considered as categories. Thus the symbolic positioning can be converted into a classification [3] problem. Classification is a well-studied part of data mining so numerous well-known classifiers could be applied for indoor positioning.

The evaluation of well-known classifier methods is usually based on the classic CRISP approach [4] which treats every misclassification equally wrong. Our previous experiments [5] shows that a more accurate classification method can predict further the symbolic positions from the original location when it is misclassified, than a less accurate classifier. In conclusion, a different approach is recommended to be used as an alternative of CRISP, which considers the topology of the building [6]. Topology of the building defines the Zones and their sizes and arrangement. The classification error should be proportional to the sizes of the Zones and the layout should have a high impact on the error values. In addition, the classification error should not necessary be symmetric due to the size differences.

A gravitational force-based approach had been presented in our previous work [7]. The approach had been tested in a simulated environment, and based on the results, the gravitational-force based approach is a promising candidate to be used instead of CRISP. However, real-life scenarios are essential for calculating classification error for symbolic indoor positioning tasks. Hence, information about a real-life indoor environment is required.

IndoorGML [8] is a standard for representing indoor spaces in an XML formatted document, where it contains the bounding three dimensional absolute coordinates for each symbolic indoor positions. Based on these documents, the sizes and the arrangement of the Zones can be derived. Thus IndoorGML standard can be used in the calculation of classification error for symbolic indoor positioning. The error calculation method is planned to be integrated into the ILONA System [9].

IndoorGML

IndoorGML [8, 10] is a standard of Open Geospatial Consortium (OGC) for specifying a data model of indoor spatial information. IndoorGML defines XML Schemas and it focuses on indoor navigation purposes. Indoor spaces are non-overlapping closed objects and they are bounded by physical or fictional boundaries. Besides the boundaries, for each space an *Envelope* is defined, which is a bounding box given by two diagonally opposite corner points, that can be used for searching.

Environment

The base environment is selected to be the three-story tall Institute of Information Science Building of the University of Miskolc. The building contains an atrium lobby, 26 offices, 15 laboratories, a lecture hall and storage rooms on nearly $3800m^2$.

The floor plan of the building can be seen in Figure 1, where ■ ✚ ★ symbols represent the ground, the first and the second floors. In the floor plan, the narrow represents the entrance of the building, and the dot is the base of the reference coordinate system. The front wall is curved, and the lobbies in the 1st and 2nd floor are bordered with curved railings, hence the shape of these rooms are not consistently rectangular cuboid or box. The statistics of the created IndoorGML Documents [11] can be seen in Table 1.

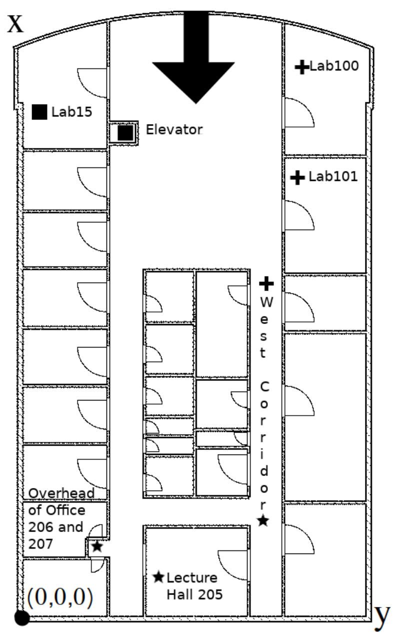

Figure 1: Floor Plan in Institute of Information Science Building

Table 1: Statistic of IndoorGML Documents of Miskolc IIS Building

Floor	Cell	Linear	Curved
Ground Floor	18	17	1
First Floor	33	32	1
Second Floor	21	20	1
Total	72	69	3

Method

The gravitational force based approach [7] should fulfill the requirements for considering the topology. The method assumes the disjunction of the Zones, and requires the determination of capacity (V) and distance (d) functions. The gravitational force [12] measures the similarity between two Zones, and it is proportional to the product of their capacity and inversely proportional to the square of their distance. The formulation of the gravitational force can be seen in Equation 22.

$$F_g(Z_i, Z_j) = \frac{V(Z_i)V(Z_j)}{d(Z_i, Z_j)^2} \tag{22}$$

While the gravitational force represents the similarity between Zones, their difference is required for error calculation. The δ function can be calculated as seen in Equation 23. It represents the difference of two zones in the $[0, 1]$ range.

$$\delta(Z_i, Z_j) = \frac{1}{1 + F_g(Z_i, Z_j)} \tag{23}$$

The classification error should be also proportional to their sizes of the zones and be asymmetric due to these size differences. The ϵ function had been introduced to measure the classification error in the $[0, 1]$ range, as it can be seen in Equation 24.

$$\epsilon(Z_i, Z_j) = \frac{V(Z_i) * \delta(Z_i, Z_j)}{V(Z_i) + V(Z_j)} \tag{24}$$

Results

For the implementation of the topology-based classification error calculation a Java application had been developed. It converts the data from the IndoorGML XML document to Java classes, both provided by IndoorGML. A zone is represented as CellSpace object, which means that each zone contains the name and the ID of the zone, the bounding and two diagonal cornerstone coordinates.

In this paper, the three-dimensional lower- and upper corner coordinates are used in distance and capacity calculation. For distance calculation the Euclidean distance had been chosen, and the distance is specified by the length of the straight line between the middle points of the two zones. To calculate the capacity of a zone, the benefit of cuboid property had been applied to calculate the volume. Based on the distance and the capacity function, the classification error can be calculated for each zone pair. Table 2 shows some examples of the classification error.

Table 2: Examples of the Topology-Based Classification Error

Actual			Predicted			Error
Name	Capacity (m^3)	Centroid	Name	Capacity (m^3)	Centroid	
Ground Floor Elevator	17.5	(40.8,8.3,4.5)	Overhead of Office 206 and 207	2.8	(5.5,6.5,4.5)	0.3641
Lab15	169.4	(45.5,2.8,1.4)	Lecture Hall 205	156.8	(3.5,14,7.6)	0.0009
Lab100	241.5	(44,8,24.8,4.5)	Lab101	172.9	(33.8,25.3,4.5)	0.0002
2nd Floor West Corridor	231	(16.5,19.8,7.6)	1st Floor West Corridor	231	(16.5,19.8,4.5)	0.0001

The overall average classification error is 0.0088 with 0.0271 standard deviation. The highest error calculated is 0.3641 in the case of *Ground Floor Elevator* actual zone is misclassified as *Overhead of Office 206 and 207*. The volume of *Ground Floor Elevator* is $17.5\ m^3$, while the *Overhead of Office 206 and 207* is $2.8\ m^3$.

The two farthest zones are *Lab15* on the front of the ground floor and *Lecture Hall 205* on the back of the second floor. The volume of these zones are 169.4 and $156.8\ m^3$. The classification error calculated in this case is 0.0009.

The *Lab100* and *Lab101* are neighbouring zones with 241.5 and $172.9\ m^3$ volume. The classification error in this case is 0.0002. The *2nd Floor West Corridor* and the *1st Floor West Corridor* zones are congruent, they only differ in the z coordinate. The misclassification in both direction result the 0.0001 value.

Discussion

The gravitational force-based method should consider the topology of the classification error calculation. It is proportional to the sizes and the arrangement of the Zones.

As it can be seen in Figure 1, the *Ground Floor Elevator* and the *Overhead of Office 206 and 207* both relatively small zones, and they are very far from each other. Hence the classification error in this case is high. Otherwise, the *Lab15* and the *Lecture Hall 205* are also very far from each other, but they are both relatively large, thus it has significantly lower error value than in the case of *Ground Floor Elevator* and

Overhead of Office 206 and 207. Therefore the method considers the size and the layout of the zones, thus the misclassification of smallest, farthest zones results the highest error values.

The misclassification of *Lab100* to *Lab101* zones has a relatively small error value. *Lab100* is larger than *Lab101*, but both are considered as relatively large zones and they are neighbouring.

The *2nd Floor West Corridor* and *1st Floor West Corridor* zones are congruent in size and location besides the z coordinate. As it would be expected, the classification error of these two zones are symmetric, and small.

The results of the classification error calculation show that the gravitational force-based approach considers the topology in classification error calculation. However, the calculated error values has a very low average, and the standard deviation should be higher, and the highest error value is lower than the half of the possible range.

Summary

To demonstrate the gravitational force-based classification error calculation approach, a Java application was developed as a proof of concept. The selected environment is the Institute of Information Science building in University of Miskolc. A model of this building had been constructed with IndoorGML standard. The IndoorGML Documents specify indoor spatial information, and they represents the indoor space as non-overlapping cells. The presented results showed that the gravitational force-based approach fulfills the criteria to consider the topology in classification error calculations. Thus it is proportional to the sizes and the layout of the zones, and it is not necessary symmetric. Therefore, the gravitational force-based approach can be an alternative to the CRISP approach in the evaluation of symbolic indoor positioning methods. Hence, the developed Java application can be used to calculate the classification error based on IndoorGML Documents.

In the future, the gravitational force-based approach will be integrated to the ILONA System. It will be also used to compare well-known classifiers for selecting the most suitable for symbolic indoor positioning purposes.

Acknowledgements

The described article was carried out as part of the EFOP-3.6.1-16-2016-00011 "Younger and Renewing University – Innovative Knowledge City – institutional development of the University of Miskolc aiming at intelligent specialisation" project implemented in the framework of the Szechenyi 2020 program. The realization of this project is supported by the European Union, co-financed by the European Social Fund.

References

[1] Hakan Koyuncu and Shuang Hua Yang. A survey of indoor positioning and object locating systems. *IJCSNS International Journal of Computer Science and Network Security*, 10(5):121–128, 2010.

[2] Hui Liu, Houshang Darabi, Pat Banerjee, and Jing Liu. Survey of wireless indoor positioning techniques and systems. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 37(6):1067–1080, 2007.

[3] Krzystof Jajuga, Andrzej Sokolowski, and Hans-Hermann Bock. *Classification, clustering, and data analysis: recent advances and applications*. Springer Science & Business Media, 2012.

[4] David Martin Powers. Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation. 2011.

[5] Judit Tamas and Zsolt Toth. Classification-based symbolic indoor positioning over the miskolc iis data-set. *Journal of Location Based Services*, 12(1):2–18, 2018.

[6] Judit Tamas and Zsolt Toth. Limitation of crisp accuracy for evaluation of room-level indoor positioning methods. In *2018 IEEE International Conference on Future IoT Technologies*, Eger, Hungary, 01.18-01.19 2018. IEEE. In press.

[7] Judit Tamas and Zsolt Toth. Topology-based classification error calculation for symbolic indoor positioning. In *Carpathian Control Conference (ICCC), 2018 19th International*, pages 643–648. IEEE, 2018.

[8] J Lee, Ki-J Li, S Zlatanova, TH Kolbe, C Nagel, and T Becker. Ogc® indoorgml. *Open Geospatial Consortium standard*, 2014.

[9] Zsolt Toth. Ilona: indoor localization and navigation system. *Journal of Location Based Services*, 10(4):285–302, 2016.

[10] Ogc indoorgml-with corrigendum. http://docs.opengeospatial.org/is/14-005r4/14-005r4.html. [Online; accessed 13-October-2017].

[11] Krisztian Ilku and Judit Tamas. Indoorgml modeling: A case study. In *Carpathian Control Conference (ICCC), 2018 19th International*, pages 633–638. IEEE, 2018.

[12] Isaac Newton. *The Principia: mathematical principles of natural philosophy*. Univ of California Press, 1999.