Modeling Moving Regions: Colorectal Cancer Case Study

Marwa Massaâbi and Jalel Akaichi

Abstract Objects in space have to be represented in order to be stored and analyzed. The three basic abstractions of spatial moving objects are moving point, line or region. The first two abstractions are highly handled. However, moving regions have always been a challenge due to their unstable shape and movement. Researchers are not giving enough attention for managing and querying this particular type of spatial data in order to solve real world problems. Motivated by this fact, we present in this paper an overview on moving regions. We survey region's modeling aspects. Then, we support our research by studying a biomedical case to highlight the importance of using moving regions. The case study illustrates the conceptual aspect of the movement of the colorectal cancer. We also use fuzzy logic thanks for its simplicity and its easy understanding. The combination offers an easier understanding for decision makers.

Keywords Moving regions · Data modeling · Fuzzy logic · Fuzzy UML

1 Introduction

A region is an object whose extent is changing over time (for example, air polluted zones, shrinking balloon, volcanic lava eruption, forests, lakes). This object has to meet certain criteria: it must have a changing extent (grow/shrink) and location (move continuously) over time. There are two types of moving regions: a simple region and a complex one (having holes inside). The difficulty in this case is when supporting and dealing with this particular type of moving objects. In other words, it is hard to capture every single movement state in every time unit and between them.

M. Massaâbi (✉)
BESTMOD Laboratory, Université de Tunis, ISG - Tunis, Tunis, Tunisia
e-mail: massaabi.marwa@gmail.com

J. Akaichi
College of Computer Science, King Khalid University, Abha, Saudi Arabia
e-mail: jalel.akaichi@kku.edu.sa

G. De Pietro et al. (eds.), *Intelligent Interactive Multimedia Systems and Services 2016*, Smart Innovation, Systems and Technologies 55,
DOI 10.1007/978-3-319-39345-2_36

Some of this region data can be imprecise or incomplete due to an inconstant sampling rate for example (every few seconds or every few days), or erroneous points (a point in the ocean). The uncertainty and the imprecision can also be caused by bad reception condition of GPS receivers. Despite the progress in this research area, the present works still require further improvements in order to make modeling moving regions an easier and more effective task. Therefore, in our work, we chose to use Fuzzy Logic to overcome this imperfection. In fact, Fuzzy Logic was introduced by Lotfi Zadeh in 1965 [18]. Fuzzy logic is an extension of the Boolean logic. It is a mathematical model distinguished by its ability to represent and process uncertain and imprecise information. In this paper, we survey the existing methods for modeling moving regions. Then, we take colorectal cancer as an example and we propose a conceptual model that represents its movement. This paper is organized as follows: Sect. 2 surveys and classifies moving regions modeling techniques. Section 3 presents a detailed case study. Finally, Sect. 4 concludes and proposes research directions.

2 Related Works

Several researchers took interest in modeling moving regions. Therefore, in this section, we recall the elaborated works in this field in order to overview its evolution over the last years. Hence, we classified the surveyed works into three categories: Crisp, Fuzzy and Conceptual modeling.

2.1 Crisp Modeling

In [16], the authors used the polygon representation to model every object moving both discretely and continuously. These objects are characterized by a dynamic shape and extent (i.e. region). They took special interest in the continuously moving regions having their data stored in a discrete way. Therefore, their work aims to estimate the in-between state of the object to restore its continuous feature. In [10], the authors presented their big data based system "Storm DB" used to answer rain fall amount queries in the continental United States. The moving regions in their case are the rain clouds. Storm DB is composed of five steps: They started by extracting the regions from radar images provided by the US national weather service. Then, they classified the regions according to their belonging to an existing storm or a new one. The next step is to calculate the duration of a point in a region in order to obtain the rainfall amount list. The fourth step aims to accelerate the algorithm used in the previous one. They concluded by the Big Data architecture. It consists of a map/reduce algorithm that supports parallelism and processes a massive amount of data. Their approach, indeed deals with moving regions. However, the modeling aspect was not their strength point since they considered regions as simple points. In [17], the authors presented a data model for route planning in the case of forest fire. They developed an application that suggests safe circuitous routes to avoid the moving flames. To do so, they used a fire simulation model. This model takes into consideration the wind properties since it influences the fire state and direction. It also simulates and gives information about the fire. This model is taken from [9]. It consists of a real time algorithm that simulates and predicts forest fire spreading. Some authors preferred well-known techniques, whether some other authors chose algebra for managing moving regions. The following works [1–3, 7] treated the same issue from almost the same point of view. Well, in [1], the authors presented their approach for modeling moving objects and discussed some issues. Among the issues, they discussed the level of abstraction suited for an algebra. For instance, a region could be represented as a polyhedral shape in the space or as a continuous function. As they consider moving objects as geometries changing over time, they defined corresponding algebraic operations. In order to specify the algebraic operations, they began by proposing an abstract algebraic data model, i.e., a model that describes data in a simple and sometimes unrealistic way. The abstract model was followed by the according discrete one. The same data structures and algorithms have already been used for spatial databases in the Rose algebra [4, 5]. In [3], the authors proposed a data model for moving objects in the two-dimensional space. They presented their system design and presented its associated operations. They specified the types' semantics and operations in the abstract model. While in [2], the authors completed the previous work [3] by presenting the discrete model. They transformed the abstract model into a discrete one for which they provided data structures and illustrating algorithms. The work elaborated in [7] combines between the three previously cited works [1–3] and enriches them. In fact, the authors developed powerful algorithms able to deal with a large set of operations and created a basis for the development of a reliable DBMS extension package for moving objects.

2.2 Fuzzy Modeling

The previously presented methods used crisp methods to deal with crisp spatial objects. However, some authors preferred to use the Fuzzy Set theory for a better representation of fuzzy spatial objects. In fact, this theory offers several properties and operations that correspond to some real life objects specificities (i.e. spatial objects with vague and imprecisely determined boundaries or interiors). In [6], the authors proposed an approach that models regions and their trajectories, and predicts their future movements as well. In this work, regions are delimited by a rectangle, also named enclosing box, to which was assigned a center point. The trajectory, in this case, is the link between the center points. After modeling regions, the authors predicted its future positions using the linear regression and the recursive motion function. They tested their application using two datasets concerning fires and storms. In his works, [11–14], Markus Schneider was interested in fuzzy spatial databases and fuzzy moving objects. We begin with [11], the researcher offered a fuzzy spatial data type for fuzzy points, fuzzy lines and fuzzy regions. He also proposed novel fuzzy spatial operations. In [12], he continued his work by proposing metric operations on fuzzy spatial objects. Then, in [13] he defined a conceptual model of fuzzy spatial objects with its implementation. The particularity of this work is the use of the discrete geometric domain (grid partition) instead of the Euclidean space. In [14], the author began by presenting a detailed overview on spatial data types and a survey concerning spatial fuzziness. Then, he proposed a fuzzy algebra to represent fuzzy data types. He focused on treating spatial vagueness.

2.3 Conceptual Modeling

Conceptual modeling is important. It offers a clearer vision and understanding of the treated issues. The conceptual diagrams and data models indeed depict and describe complex objects and their relations. In [8], the authors proposed a UML data model that combines fuzzy sets and possibility theory to overcome data fuzziness and manage complex objects. They presented their contributions at various levels: fuzzy class, fuzzy generalization, fuzzy aggregation, fuzzy association and fuzzy dependency. Then, they presented the corresponding fuzzy relational schema with the transformation steps. Another work, [15], in which the authors presented a conceptual model for fuzzy object-oriented databases using UML. Their model targets imprecise data, which they called fuzzy data. In fact, they defined fuzzy objects, classes and the relations between them, together with the corresponding notations.

2.4 Recapitulation

To recapitulate, we present the following table (Table 1).

Table 1 classifies and resumes the previously cited methods for modeling moving regions. The works are cited in the table in the chronological order.

According to the presented state of the art, several techniques were used to model moving regions. Among the years, researchers took advantage of fuzzy set theory, algebra, big data, etc. Each and every one of these works has its advantages and drawbacks. Therefore, in this paper, we introduce a UML conceptual model to represent the movement of cancer cells. In fact, constructing an accurate corresponding model allows capturing the basic concepts and relations between them. Building the model is in other way describing the problem in order to understand and solve it.

Table 1 Classification of the presented methods

Cited works	Crisp modeling	Algebraic modeling	Fuzzy modeling	Conceptual modeling
[1]		Algebra		
[11]			Fuzzy Sets	
[3]		Algebra		
[12]			Fuzzy Sets	
[16]	Polygon representation			
[7]		Algebra		
[13]			Fuzzy sets	
[14]			Fuzzy sets	
[6]	Enclosing box technique			
[8]				UML
[10]	Big data approach			
[15]				Fuzzy UML
[17]	Simulation algorithm			

3 Case Study

Analyzing moving objects trajectories is a very interesting and general field. One of its major advantages is that it can be applied in several domains. In this case, we are interested in the medical field: observing illnesses, examining their behavior from a patient to another, and why not, predicting their effects on different patients too. Well, diseases move in regions, like fires, they neither have an exact shape nor follow a definite path. For this reason, we chose one specific disease (cancer). In fact, there are several types of cancers, but we specifically chose colorectal cancer. It starts in the colon then travels to the whole body in evolving speed. It moves in a group of countless microscopic cells characterized by irregular shape, size and borders. In order to clarify the fuzzy aspect in the transformation phase, we propose a conceptual model that describes the moving cancer cells. The proposed conceptual model is characterized by the fuzzy class. A fuzzy class is a class that describes the fuzzy object with at least one fuzzy attribute. A fuzzy attribute has linguistic values instead of numerical ones. The moving object is represented by a fuzzy class named ‘Cancer’ (Fig. 1).

It is defined by the following attributes: Speed, Size and Age. ‘Speed’ describes the tumor’s propagation speed. The tumor propagation speed is very slow in the first fifteen or twenty years which, unfortunately, makes the early detection very difficult. In the following ten years the speed increases and becomes noticeable. In the final stage, i.e., the last two years, the tumor spreads widely in a high speed. ‘Size’ provides the cancer cells size which tells about the gravity of the situation. And ‘Age’

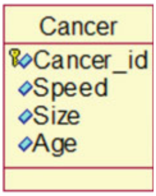

Fig. 1 The class 'Cancer'

represents the diagnosed age of the illness. Among these attributes, only two are represented fuzzily: 'Speed' and 'Size' are better identified using linguistic values. The set of linguistic values defining the attribute 'Speed' is {low, medium, high}. The set of linguistic values defining the attribute 'Size' is {tiny, small, big}. At this level, the membership function should be calculated for each linguistic value.

The membership function describing the fuzzy attribute 'Speed' is given below:

$$\mu_{low}(Speed) = \begin{cases} 1 & \text{if } Speed \leqslant 20 \\ 0 & \text{if } Speed \geqslant 25 \\ \frac{25-Speed}{25-20} & \text{if } 20 < Speed < 25 \end{cases}$$

$$\mu_{medium}(Speed) = \begin{cases} 0 & \text{if } Speed \leqslant 20 \\ 0 & \text{if } Speed \geqslant 40 \\ 1 & \text{if } 35 \leqslant Speed \leqslant 25 \\ \frac{Speed-20}{25-20} & \text{if } 20 < Speed \leqslant 25 \\ \frac{40-Speed}{40-35} & \text{if } 40 \leqslant Speed < 35 \end{cases}$$

$$\mu_{high}(Speed) = \begin{cases} 0 & \text{if } Speed \leqslant 35 \\ 1 & \text{if } Speed \geqslant 40 \\ \frac{Speed-35}{40-35} & \text{if } 35 < Speed < 40 \end{cases}$$

The membership function of the fuzzy attribute 'Speed' is graphically represented by Fig. 2. The curve defines to each point a membership value between 0 and 1.

The membership function describing the fuzzy attribute 'Size' is given:

$$\mu_{tiny}(Size) = \begin{cases} 1 & \text{if } Size \leqslant 0.5 \\ 0 & \text{if } Size \geqslant 1 \\ \frac{1-Size}{1-0.5} & \text{if } 0.5 < Size < 1 \end{cases}$$

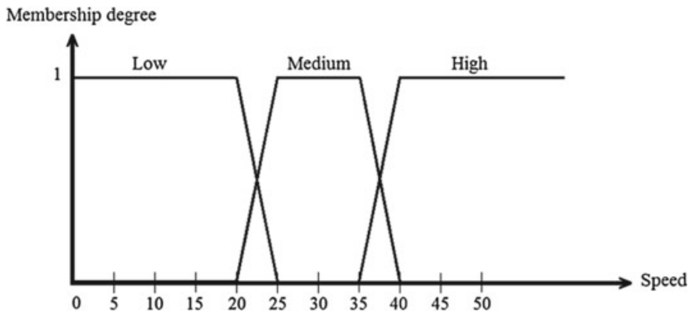

Fig. 2 The membership function of the attribute 'Speed'

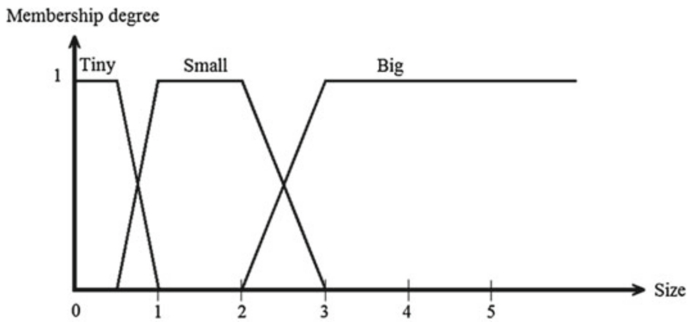

Fig. 3 The membership function of the attribute 'Size'

$$\mu_{small}(Size) = \begin{cases} 0 & \text{if } Size \leqslant 0.5 \\ 0 & \text{if } Size \geqslant 3 \\ 1 & \text{if } 1 \leqslant Size \leqslant 2 \\ \frac{Size-0.5}{1-0.5} & \text{if } 0.5 < Size \leqslant 1 \\ \frac{4-Size}{4-3} & \text{if } 4 \leqslant Size < 3 \end{cases}$$

$$\mu_{big}(Size) = \begin{cases} 0 & \text{if } Size \leqslant 2 \\ 1 & \text{if } Size \geqslant 3 \\ \frac{Size-2}{3-2} & \text{if } 2 < Size < 3 \end{cases}$$

The membership function of the fuzzy attribute 'Size' is graphically represented by Fig. 3. The curve defines to each point a membership value between 0 and 1.

After defining the attributes and their membership functions, we created the fuzzy general conceptual model of the cancer's movement using a Unified Modeling Language diagram (Fig. 4).

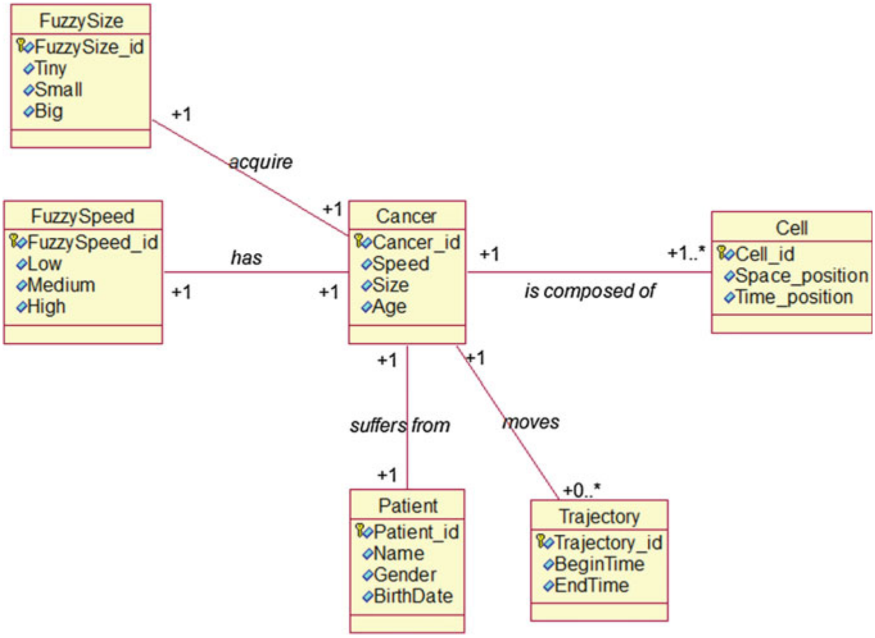

Fig. 4 The fuzzy conceptual model

The model describes the interaction between the tumor itself and the surrounding objects that could affect it. The class 'Cancer' is related to three classes: 'Patient' which describes the patient having the tumor, 'Trajectory' which describes the trajectory of the tumor in the human body, and more importantly 'Cell'. Actually, the 'Cell' class is the one that identifies the cancer as a region. It is in this case represented by a set of points not one single point. This representation reflects in a more accurate way the nature of this moving object. In addition to the diagram classes, we added two fuzzy classes: 'FuzzySize' and 'FuzzySpeed'. In fact, every linguistic attribute of the class 'Cancer' is represented in the conceptual model by a fuzzy class. Each fuzzy class is characterized by its primary key and its linguistic values as attributes.

Once the conceptual model is done, we moved to next one. The logical model of the cancer's movement using the snowflake schema is presented by Fig. 5. It describes the relations between the tables and it is inspired by the conceptual model. Both model were elaborated by Rational Rose which is a software published by the company Rational, and acquired by IBM, to create and edit UML diagrams.

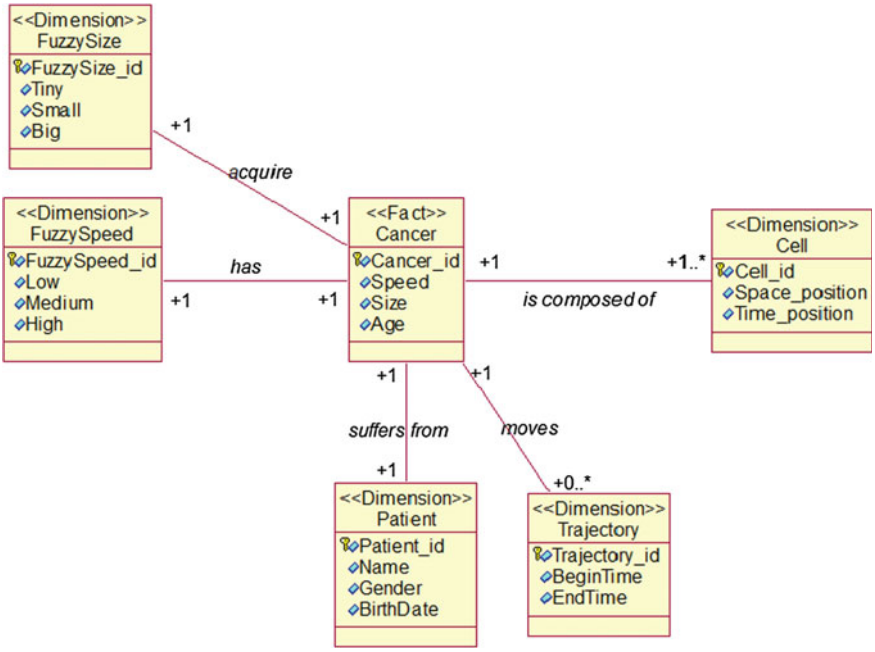

Fig. 5 The fuzzy logical model

4 Conclusions and Future Works

In the recent years, researchers around the world are more and more interested in studying moving object data (points, lines and regions). Moving regions, in particular, are characterized by their continuously changing shape and path together with some possible holes inside sometimes. The regions' specifications inspired us to use fuzzy logic to have a better representation, storage and analysis. In this paper, we surveyed the issue of modeling moving regions. We also reviewed limitations of the current systems. Unfortunately, the current generation of moving regions modeling systems still requires further improvements to obtain better results. Hence, we began by presenting a moving cancer conceptual model. This model is a first step to a complete model that deals with moving regions. We presented the integration of fuzzy logic into the regions' modeling using fuzzy membership functions, also, by providing both conceptual and logical fuzzy models. The next step will be to enrich our model by a complete framework dealing with moving regions' data. The framework will be able to analyze and predict the behavior of this particular kind of data in order to support the decision making process.

References

1. Erwig, M., Güting, R.H., Schneider, M., Vazirgiannis, M.: Spatio-temporal data types: an approach to modeling and querying moving objects in databases. Geoinformatica **3**(3), 269–296 (1999). http://dx.doi.org/10.1023/A:1009805532638
2. Forlizzi, L., Güting, R.H., Nardelli, E., Schneider, M.: A data model and data structures for moving objects databases, vol. 29. ACM (2000)
3. Güting, R.H., Böhlen, M.H., Erwig, M., Jensen, C.S., Lorentzos, N.A., Schneider, M., Vazirgiannis, M.: A foundation for representing and querying moving objects. ACM Trans. Database Syst. **25**(1), 1–42 (2000). http://doi.acm.org/10.1145/352958.352963
4. Güting, R.H., De Ridder, T., Schneider, M.: Implementation of the rose algebra: efficient algorithms for realm-based spatial data types. In: Advances in Spatial Databases, pp. 216–239. Springer (1995)
5. Güting, R.H., Schneider, M.: Realm-based spatial data types: the rose algebra. VLDB J. **4**(2), 243–286 (1995)
6. Junghans, C., Gertz, M.: Modeling and prediction of moving region trajectories. In: Proceedings of the ACM SIGSPATIAL International Workshop on GeoStreaming, pp. 23–30. ACM (2010)
7. Lema, J.A.C., Forlizzi, L., Güting, R.H., Nardelli, E., Schneider, M.: Algorithms for moving objects databases. Comput. J. **46**(6), 680–712 (2003)
8. Ma, Z., Zhang, F., Yan, L.: Fuzzy information modeling in uml class diagram and relational database models. Appl. Soft Comput. **11**(6), 4236–4245 (2011)
9. Moreno, A., Segura, Á., Korchi, A., Posada, J., Otaegui, O.: Interactive urban and forest fire simulation with extinguishment support. In: Advances in 3D Geo-Information Sciences, pp. 131–148. Springer (2011)
10. Olsen, B., McKenney, M.: Storm system database: a big data approach to moving object databases. In: 2013 Fourth International Conference on Computing for Geospatial Research and Application (COM. Geo), pp. 142–143. IEEE (2013)
11. Schneider, M.: Uncertainty management for spatial datain databases: fuzzy spatial data types. In: Advances in Spatial Databases, pp. 330–351. Springer (1999)
12. Schneider, M.: Metric operations on fuzzy spatial objects in databases. In: Proceedings of the 8th ACM international symposium on Advances in geographic information systems, pp. 21–26. ACM (2000)
13. Schneider, M.: Design and implementation of finite resolution crisp and fuzzy spatial objects. Data Knowl. Eng. **44**(1), 81–108 (2003)
14. Schneider, M.: Fuzzy spatial data types for spatial uncertainty management in databases. In: Handbook of Research on Fuzzy Information Processing in Databases, vol. 2, pp. 490–515 (2008)
15. Singh, S., Agarwal, K., Ahmad, J.: Conceptual modeling in fuzzy object-oriented databases using unified modeling language. Int. J. Latest Res. Sci. Technol. **3**, 174–178 (2014)
16. Tøssebro, E., Güting, R.H.: Creating representations for continuously moving regions from observations. In: Advances in Spatial and Temporal Databases, pp. 321–344. Springer (2001)
17. Wang, Z., Zlatanova, S., Moreno, A., van Oosterom, P., Toro, C.: A data model for route planning in the case of forest fires. Comput. Geosci. **68**, 1–10 (2014)
18. Zadeh, L.A.: Fuzzy sets. Inf. Control **8**(3), 338–353 (1965)