Journal of Theoretical and Applied Information Technology
28th February 2018. Vol.96. No 4

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

A BAYESIAN NETWORK APPROACH TO MINE SPATIAL DATA CUBE

[1]MIDOUN MOHAMMED, [2]BELBACHIR HAFIDA

[1]Dept. Informatique Université des Sciences et de la Technologie Mohamed Boudiaf Oran, Oran Algeria

[2] Dept. Informatique Université des Sciences et de la Technologie Mohamed Boudiaf Oran, Oran Algeria

E-mail : [1] miodun.mohamed@univ-usto.dz, [2]h_belbach@yahoo.fr

ABSTRACT

Spatial data mining is an extension of data mining that considers the interactions in space. It involves various techniques and methods in various areas of research. It takes into account the specificities of spatial information such as spatial relationships that can be topological, metric or directional. These relationships are implicit and difficult to represent. A Bayesian network is a graphical model that encodes causal probabilistic relationships among variables of interest, which has a powerful ability for representing and reasoning and provides an effective way to spatial data mining. Moreover, spatial data cubes allow storage and exploration of spatial data. They support spatial, non-spatial and mixed dimensions. A spatial dimension may contain vector and raster data. The spatial hierarchies can represent topological relationships between spatial objects. We propose to use Bayesian networks for knowledge discovery in spatial data cubes. The goal of our approach is first to consider spatial relationships in the data mining process, and secondly to benefit from the strength of the data warehouses to apply spatial data mining on different aggregation levels according to the topological relations between spatial data. In this article, we give a state of the art on spatial data mining and propose a framework for data mining in spatial data cubes, using Bayesian networks. We show in the proposed case study that our approach confirms the results observed in the field and it is an important way to take into account the specificities of spatial data in the spatial data mining process.

Keywords: *F Spatial data mining, Bayesian networks, Spatial data cube, Spatial aggregation, Spatial Analysis*

1. INTRODUCTION

The large databases currently available have a strong spatial component.

The quantity of these spatial or georeferenced data will continue to increase during the 21st century. Examples include terabytes of georeferenced data generated daily by Earth observation satellites or climate and environmental databases.

One of the challenges of research is to analyze this data and discover new knowledge in the form of models and relationships.

Spatial analysis methods traditionally used by Geographic Information Systems (GIS) for spatial data mining were developed at a time when data sets were small.

They are confirmatory, require initial assumptions, and focus on obtaining scarce information from small data sets.

They are not meant to discover unexpected new patterns, trends, and relationships that can be hidden in very large and heterogeneous sets of spatial data.

Spatial data mining is an extension of data mining that considers the interactions in space. It takes into account the specificities of spatial information, such as spatial relationships and spatial dependence.

Knowledge discovery in spatial databases differs from knowledge discovery in traditional relational databases.

Indeed, the representation of spatial information requires an implicit topological and geometric measurement framework that affects the patterns that can be extracted.

In addition, spatial data is also spatially dependent, which means that similar elements cluster in space.

Added to this, the spatial data are poorly structured and scale dependent.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Finally, we must consider the spatial heterogeneity of the data, which means that spatial data are not homogeneously distributed in geographical space.

These properties make conventional data mining algorithms, which assume that data is generated independently and distributed identically in space, are not appropriate for spatial data.

Many works have been proposed, they concern: association rules, clustering, classification methods or Bayesian networks.

A Bayesian network is a graphical model that encodes causal probabilistic relationships among variables of interest, which has a powerful ability for representing and reasoning and provides an effective way for spatial data mining.

On the other hand, special data cubes are cubes that contain dimensions or facts which are spatially referenced and can be represented on maps [1]. They allow the storage and exploration of spatial data. They support spatial, non-spatial and mixed dimensions. A spatial dimension can contain vector data (geometry) or raster data (image).

The nature of spatial information generates a set of problems of incompatibility with the principles of data mining.

First, the spatial data is linked, while the methods of datamining consider that the data are independent. On the other hand, the spatial relationships are implied and are seldom stored in databases.

The spatial relations are multiple, they may be topological (adjacency, intersection...) or metric (distance) and the analysis can be mono or multi Thematic. This makes it difficult to choose the correct spatial relationship.

To represent the spatial relationships in relational databases, we can use the spatial joint index or contiguity matrix.

Another approach is to model the spatial information into spatial data cubes. A spatial data cube is an ideal environment for data mining, it allows analysis and spatial queries on several levels of spatial aggregation.

Several works on data mining on spatial data cubes were made. However, few studies have applied Bayesian networks on spatial data cubes. This, due to the complexity of spatial data sets.

Our major contribution is to propose a platform for the application of Bayesian networks on spatial data cubes for data mining purposes.

To represent spatial relationships, we use a spatial hierarchy of vector layers that will respect the topological relationships between spatial objects.

The spatial aggregation will be used to calculate the measures and then apply data mining on different levels of the spatial hierarchy.

The main interest of our contribution is to use Bayesian networks to apply spatial data mining on different levels of aggregation of spatial hierarchy.

This by considering the spatial relationships, including topological relationships implicitly stored in a spatial data cube. We will use spatial analysis to confirm the validation of our approach and view the results on a map.

Without considering the specifics of spatial data, data mining methods cannot discover hidden relationships.

Our approach allows us to take into account topological relationships between spatial data and to explore databases at different scales of analysis. This facilitates the use of Bayesian networks and makes them applicable for the exploration of spatial data.

In the next section we give an overview of some existing works pertaining to spatial data mining and spatial data cubes.

Then we define our approach and we propose a framework of spatial data mining based on Bayesian networks.

The results and evaluation of our approach will be discussed in experiments section. Finally, we end this paper with some conclusions.

2. RELATED WORKS

Spatial data mining is the application of data mining techniques to spatial data. It can be defined as the discovery of interesting, implicit and previously unknown knowledge from large spatial data bases [2].

The main objective of the spatial data mining is to discover relationship and characteristics that may exist implicitly in spatial databases. It has been used in various fields like remote sensing, medical imagery and Visual data mining.

Spatial Data Mining extends relational data mining with respect to special features of spatial data, like

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

mutual influence of neighboring objects by certain factors (topology, distance, direction).

Extracting interesting and useful patterns from spatial datasets is more difficult than extracting the corresponding patterns from traditional numeric and categorical data due to the complexity of spatial data types, spatial relationships, and spatial autocorrelation.

Many works have been proposed for spatial data mining, they relate to the various tasks of data mining, such as classification [3,4], association rules [5], or clustering. [6].

Moreover, the application of Bayesian Networks for spatial data mining and knowledge discovery was introduced by [7]. Bayesian networks provide a coherent framework of representation and reasoning for spatial problems.

The process of spatial data mining based on Bayesian networks includes two parts, one is structure learning, and the other is learning the parameters of the network.

Many research has studied on structure learning [8,9], and many other research has studied the algorithms and approaches to learning the parameters [10,11].

As for a spatial data mining method, Bayesian networks can be used for spatial knowledge representation, spatial classification, spatial clustering, and spatial prediction [7].

Several studies have been conducted: [12] used Bayesian network classifiers for mineral potential mapping, [13] developed an algorithm that can be applied to large trajectory collections, [14] proposed spatial Bayesian learning algorithms for geographic information retrieval, and [15] proposed a Bayesian method for assessing vulnerability to natural disasters to catastrophic risk.

Data warehouses are databases of information dedicated to the analysis and decision making [16]. A data is a subject-oriented, integrated, time-variant and non-volatile [17].

Spatial data warehouse is data warehouse where some dimension members or some facts are spatially referenced and can be represented on a map.

Spatial data warehouses contain geographic data, for example, satellite images, and aerial in addition to non-spatial data.

A number of studies have been conducted for spatial data mining in spatial data cubes. They relate in particular to the use of association rules, classification methods, and exploitation of raster databases (Image). [18,19]

The main difficulty in spatial data mining is the recognition of spatial relationships in databases. These spatial relationships are implicit and difficult to be represented. Several solutions have been proposed:

Indeed, the spatial relationships between objects in a spatial framework are often modeled by a contiguity matrix [20]. A contiguity matrix can be representing a neighborhood relationship defined using the Euclidean distance or contiguity.

Another solution proposed by Valduriez [21] is to add a joint index to speed up the joints as part of a relational database. The extension to spatial data has been proposed by Zeitouni et al. in [22].

This extension consists of adding a third attribute that represents the spatial relationship between two objects More models of spatial relationships using hypergraphs are available in the literature.

[23] propose to model topological relationships through spatial hierarchies of spatial data cube. They define the different types of spatial hierarchies.

In addition, they classify topological relationships between hierarchical levels according to the procedures required for ensuring correct measure aggregation.

A spatial data cube can include numerical measures and spatial measures and pointers to spatial objects at different levels of aggregation. Aggregation of spatial objects is not easy; it requires the use of a spatial hierarchy.

Few studies have applied Bayesian networks on spatial data cubes. This research doesn't consider the specificities of spatial data.

The application of Bayesian networks does not take into account spatial and topological relationships

Journal of Theoretical and Applied Information Technology
28th February 2018. Vol.96. No 4

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

and considers a single level of detail, while the analysis and exploration of spatial data has to be done on several levels of scales.

Our contribution is to provide a methodology for the application of Bayesian networks on spatial data cubes.
With the aggregation of spatial and non-spatial measures, our work allows to taken in account the spatial relationships, including topological relationships between different objects, and perform a knowledge discovery in various aggregation levels of a spatial hierarchy.

3. OUR APPROACH

Figure 1 shows the approach we propose for the application of Bayesian networks on spatially referenced data. We use GIS data that we store in a database.

These data are then integrated in a cube of spatial data, where the measurements are aggregated according to the different levels of the spatial hierarchy which corresponds to the topological relations of the spatial objects.

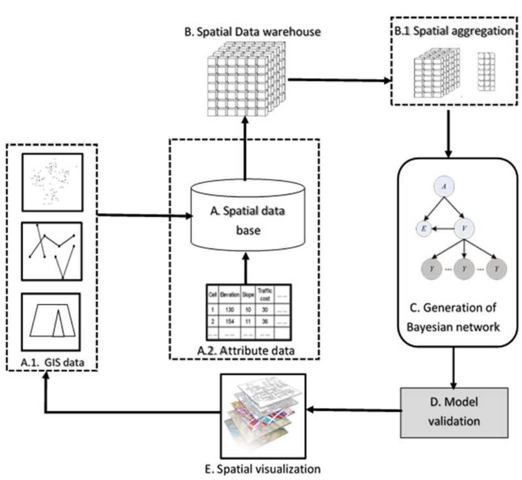

Figure 1: Spatial data mining procedure using Bayesian networks

Then we apply the Bayesian networks for data mining purposes to predict the progress of constructions in a housing program.

The results obtained are then validated by comparing them with the results observed in the field. In the following we describe in more detail the main steps of our approach.

3.1 Dataset presentation

In a Geographic Information System (GIS), there are vector data (geometric: point, line, polygon) and raster data (pixels).

Our approach is based on GIS vector data. We get the data from multiple heterogeneous sources of spatial data (vector) and non-spatial data (attribute).

Then, we apply the pre-processing steps, such as, converting the vector data or adding spatial projection. We perform these pre-processing in a GIS environment, and then we build our spatial database.

In the spatial data mining environment, we use not only GIS to manage and visualize spatial data, but also as a means of calculating spatial measures using spatial analysis techniques.

3.2 Spatial data cube and spatial aggregation

In spatial datawarehouse, spatial information can be integrated as dimensions or measures. Spatial data cubes are cubes for which members of dimensions or facts (via spatial measures) are spatially referenced and can be represented on maps [1].

There are two types of spatial data cubes, vector cubes and raster cubes. They contain at least one dimension where some or all members are geometric.

In a data cube, data is organized in dimensions which describe in a natural way most of the attributes associated with the data of interest [24]. The dimensions are in turn organized into hierarchies, with data aggregated at each level.

As for the dimension hierarchies, topological relationships have hierarchical structures. these relations correspond to the hierarchical semantic relationships between spatial objects.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Therefore, our approach is based on the use of these topological relationships to add levels to the spatial hierarchy of our spatial data cube.

The measures of the fact table will be aggregated and calculated according to each level of aggregation of the spatial dimension. They will represent the variables on which we apply Bayesian networks.

3.3 Bayesian Network

Bayesian networks are graphical models for defining probabilistic relationships between variables.

An advantage of Bayesian networks is that they capture knowledge in a form people can understand intuitively, and which allows a clear visualization of the relationships involved.

Bayesian networks used a directed acyclic graph (DAG) to represent assertions of conditional. The nodes in the graph represent the variables and the directed arcs define the conditional relationships.

The advantages of directed graphic models over undirected models are the notion of causality. Causality indicates that if an arc is directed from A to B in the network, then A causes B. Bayes' theorem is used to calculate causal inference about the variables. Bayes' theorem states:

$$P(B_i|A)=(P(A|X_i)\,P(B_i))/(P(A)) \qquad (1)$$

where ($i= 1, 2\ldots, r$)

Though, the construction of Bayesian networks is a hard task and the number of possible structures and the number of parameters for those structures can be huge.

Learning a Bayesian network from data involves two tasks: Estimating the probabilities for the conditional probability tables (learning parameters) and deriving the structure of the network.

The process of building the Bayesian network consists of three steps: variables definition, structure learning, and parameter estimation.

Variable definition. Defines the relevant variables and the relationship between them.

Structure learning. Determine the directions of all edges based on prior knowledge and the given data set. Structure learning of Bayesian networks is the key step to perform reasoning and predicting.

Parameter estimation. It refers to define the conditional probabilities of the relationships. This step defines the conditional probabilities associated with each node.

As for a spatial data mining method, Bayesian networks can be used for spatial knowledge representation, spatial classification, spatial clustering, and spatial prediction.

Bayesian networks involves different search algorithms for constructing the network topology. The heuristic algorithms include K2, DAG, Hill Climbing (HC), and TAN (Tree Augmented Naive (Bayes)).

In our approach, we use Bayesian networks on measures of spatial data cube. The measures are calculated using aggregated levels of spatial hierarchy. Then, they will be discretized, and several Bayesian networks can be built from these measures.

3.4 Evaluation and model validation

Once the Bayesian networks are built, each network must be evaluated. For this purpose, we compare the results obtained with real results observed in the field.

The accuracy of the evaluation and the calculation of the Kappa index will allow us to evaluate the results obtained by our approach based on Bayesian networks.

3.5 Spatial analysis and data visualization

Spatial analysis is a set of methods and tools which enable to understand, evaluate and interpret the spatial distribution of phenomena in order to discover and / or highlight the general rules of organization of space [25].

Spatial analysis can be applied to the interrogation of thematic, geometric and topological components of the spatial information contained in the GIS.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Once the Bayesian model generated and validated, we integrate the parametric data of the Bayesian model, in the thematic GIS layers.

This will be performing spatial analyzes to visualize the results on a map and will validate the analysis of Bayesian network generated and visualize spatially the results obtained by comparing them with the real situation on the field.

4. EXPERIMENTS

This section describes the experiment conducted to evaluate the proposed approach. We apply Bayesian networks for Urban Planning in order to predict the progress of housing construction programs in Algeria. We use real data stored in a spatial data cube.

4.1 Experimentation environment

We applied our approaches under Windows environment, with SQL Server 2012 as database management system and ArcGIS 10.3 desktop as GIS. ArcGIS is a suite of geographic information software or GIS software developed by the American company Esri (Environmental Systems Research Institute, Inc.).

4.2 Dataset presentation

For the purposes of our experiments, we used a database comprised of vector GIS data prepared in ArcGIS environment, and non-spatial data from progress reports on housing construction programs in Algeria. This database includes 422 buildings, 143 parcels and 96 islets. Figure 2 gives a presentation of the study area.

Figure 2: Presentation of the study area

In this step, we prepare the data needed for the spatial data mining process. For this, we proceed to the conversion and integration of vector data, as well as the addition of spatial projections. Then we model and construct the GIS of the study area.

We perform these pretreatments in a GIS environment, and then build our spatial database.

These data were processed and integrated in spatial data cube modeled in snowflake schema as shown in Figure 3. This cube has four dimensions: Report, Date, Phase, and Buildings, and one spatial hierarchy: Building / Parcels/ Islets / Area / Land-use planning. This hierarchy represents the topological relationship between spatial objects classes.

There are 21 100 measures in the fact table of spatial data cube. They provide information on the progress of construction of buildings.

The measures are aggregated and calculated according to the spatial hierarchy, so as to apply Bayesian networks on different levels of aggregation of spatial data cube.

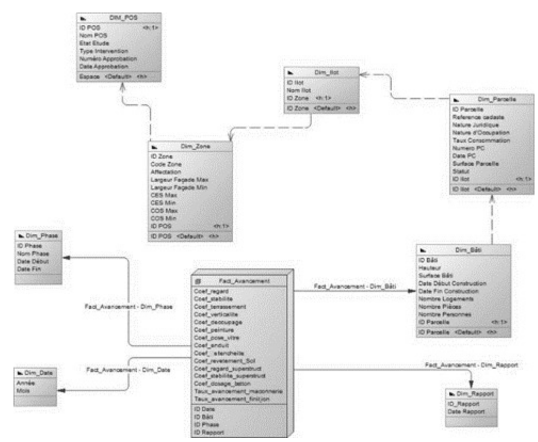

Figure 3: Representation of spatial data cube in the form of snowflake schema

4.3 Bayesian networks

In our implementation we use the K2 algorithm to calculate the Bayesian network structure. We begin by defining variables field. The variables are described as following: Earthwork, Boundary marking, Concrete dosage, Verification of verticality, Verification of stability, Floor Coating, Partitioning, Coating, Window installation, Waterproofing, Painting.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

The values of its variables are represented in the form of measures in the spatial data cube. We perform the discretization and aggregation of these measures according to the levels of the spatial hierarchy.

After defining the domain variables and data preparation, we can obtain the structure of the Bayesian network and then we should compute the conditional probabilities of the relationship.

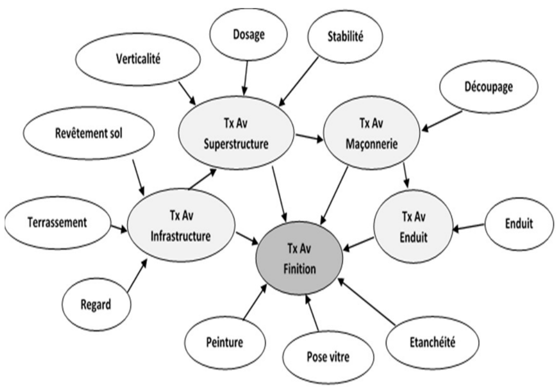

Figure 4: Bayesian network applied for monitoring the building construction process

Figure 4 shows the structure of the Bayesian network applied for monitoring the building construction process.

We apply Bayesian networks for each level of aggregation of spatial hierarchy (Buildings, parcels, islets).

Table 1: Different states of the target variable

Classe	Advancement Factor (%)	Probabilities
I	0– 10	>0.1
II	11 – 25	0.11 – 0.25
III	26 – 50	0.26 – 0.50
IV	51 – 75	0.51 – 0.75
V	76 – 99	0.76 – 0.99
VI	100	1

The Parametric results of the Bayesian network distribute the buildings, parcels and islets in six classes we have defined to represent the different stages of progress of the construction process. As shown in Table I.

4.4 Evaluation and model validation

We selected 255 cases for testing the validity of the model. Table II shows a confusion matrix. It shows the results of the experiment.

Table 2: Confusion matrix of Bayesian network

		Observed results						
		I	*II*	*III*	*IV*	*V*	*VI*	*Total*
Evaluated results	*I*	*17*	*0*	*2*	*1*	*2*	*1*	*23*
	II	*2*	*20*	*3*	*1*	*0*	*1*	*27*
	III	*0*	*1*	*36*	*3*	*0*	*2*	*42*
	IV	*1*	*0*	*4*	*57*	*3*	*1*	*66*
	V	*1*	*0*	*2*	*1*	*30*	*3*	*37*
	VI	*0*	*2*	*0*	*1*	*1*	*56*	*60*
Total		21	23	47	64	36	64	255
Evaluation accuracy /%		80,9	86,9	76,5	89,0	83,3	87,5	

We compare the results obtained with Real Results observed in the field.

The accuracy of the evaluation is 84.7% and the Kappa index is 0848. The experimental results validate the proposed approach for the spatial data mining.

4.5 Spatial analysis and data visualization

After validating the model, we integrate the parametric data of the Bayesian model in the GIS thematic layers: Buildings, parcels and islets.

Journal of Theoretical and Applied Information Technology
28th February 2018. Vol.96. No 4

ISSN: **1992-8645** www.jatit.org E-ISSN: **1817-3195**

This allows us to perform spatial analysis, visualize the results on a map and compare them with the results observed in the field.

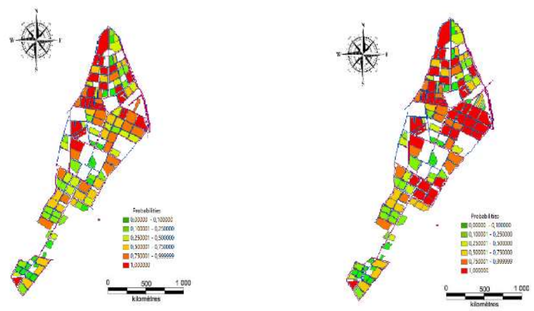

Figure 5: Spatial analysis of the distribution of buildings into classes according to the rate of progress of construction and comparison of results with observed results

Figure 5, Figure 6 and Figure 7 shows the spatial distribution of buildings, parcels and islets. They are divided on the six classes that represent the stages of completion of the construction process.

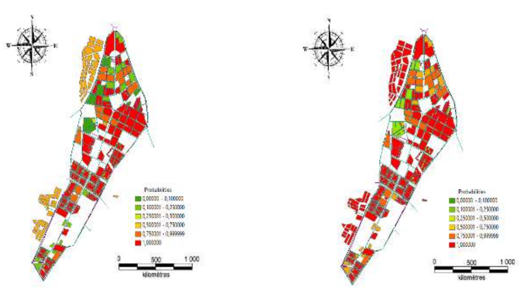

Figure 6: Spatial analysis of the distribution of parcels into classes according to the rate of progress of construction and comparison of results with observed results

In these figures we visually compare the results obtained with the results observed in the field on several levels of aggregations which correspond to the topological relations of the spatial objects.

Spatial analyzes have enabled us to spatially visualize results and confirm the validation of our approach.

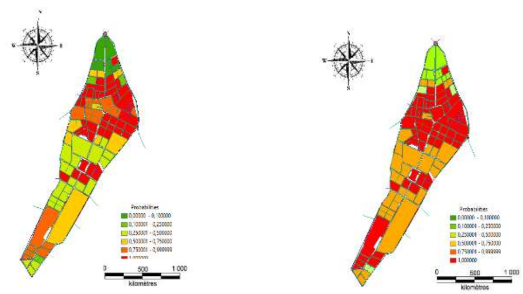

Figure 7: Spatial analysis of the distribution of islets into classes according to the rate of progress of construction and comparison of results with observed results

4.6 Discussion

A s shown in Table I, the estimation accuracy was 84.7% and the Kappa index was 0.84 which is considered as a good result for prediction.

On the other hand, using spatial analysis, the comparison of the results obtained with the results observed in the field validates our approach by a visual analysis.

We can conclude that the experimental results thus validate the feasibility of the proposed approach for knowledge discovery in spatial data.

Moreover, the application of data mining on a spatial data cube allows a knowledge discovery about the different levels of aggregation of spatial hierarchy.

Our approach allows not only to predict the construction progress of each building, but also the overall assessment of the construction process on the different islets and parcels of the study area.

Another advantage of our method is to use spatial analysis and GIS to visualize, validate and locate the results on a map.

Our approach facilitates the use of Bayesian networks for spatial data exploration. This makes it possible to generalize the use of Bayesian networks to other spatial application. We can therefore say that our approach is a good way for spatial data mining in spatial data cubes.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

5. CONCLUSION

Spatial data mining is an extension of data mining that takes into account the spatial relationships. Spatial relationships are difficult to be represented in databases.

Few studies use Bayesian networks to predict spatial phenomena, because spatial relationships cannot be explored if they are not materialized.

Our approach pre-processes spatial data by materializing spatial relationships, allowing Bayesian networks to explore spatial data. The use of spatial data cubes makes it possible to explore topological relationships in the spatial hierarchy, which is not the case with other methods of representing spatial relationships.

The case study, give results that show that Bayesian networks, despite the fact that are not adapted to spatial data, give good results of prediction if our approach is used.

In this article, we first explain a certain number of concepts related to data mining and spatial data cubes. Then, we proposed a framework for data mining in spatial data cubes using Bayesian networks.

Furthermore, we showed a case study and used the experimental data to validate the applicability of Bayesian networks for spatial data mining.

Consequently, we consider our approach as a good way to explore the spatial data.

The first interest of our approach is that it takes into consideration the spatial relationships including topological relationships.

In addition, it allows knowledge discovery about the different levels of aggregation of spatial hierarchy. Another advantage of our method is to use spatial analysis and GIS to evaluate, visualize and locate the results on a map.

However, it is necessary to apply our approach to a larger database to better judge its effectiveness. On the other hand, we also plan to apply our approach to raster data.

In conclusion, our study presents multiple perspectives, such as the development of a decision support tool that combines spatial analysis and Bayesian networks, or the development of new algorithms for Bayesian networks taking into account the spatial relationships in the process of knowledge discovery.

REFRENCES:

[1] BÉDARD, Y., MERRETT, T., HAN, J. (2008), «Fundamentals of Spatial Data Warehousing for Geographic Knowledge Discovery», In: Geographic Data Mining and Knowledge Discovery, 2e edition, chap.3, Taylor & Francis.

[2] Zeitouni K, Fouille de données spatiales, Revue internationale de géomatique n° 4/99, Numéro spécial, Edition Hermès Sciences, Avril 2000.

[3] C. E. Warrender and M. F. Augusteijn. Fusion of image classifications using Bayesian techniques with Markov rand fields. International Journal of Remote Sensing, 20(10):1987–2002, 1999.

[4] ESTER M., KRIEGEL H.-P., SANDER J., "Spatial Data Mining: A Database Approach", Proc. 5th Symp. on Spatial Databases, Berlin, Germany, 1997.

[5] M. Kamber, J. Han, and J. Y. Chiang. Metarule-guided mining of multi-dimensional association rules using data cubes. KDD'97, 207-210, Newport Beach, California. 1997

[6] Han, J.,M. Kamber, and A. K. H. Tung, "Spatial Clustering Methods in Data Mining: A Survey", H. Miller and J. Han (eds.), Geographic Data Mining and Knowledge Discovery, Taylor and Francis, 2001.

[7] J. Huang and Y. Yuan, "Construction and application of bayesian network model for spatial data mining," in Control and Automation, 2007. ICCA 2007. IEEE International Con- ference on. Institute of Electrical and Electronics Engineers (IEEE), 2007, pp. 2802–2805

[8] W. Lam, F. Bacchus, "Learning Bayesian Belief Networks: An approach based on the MDL Principle," Computational Intelligence. 10, 269-293(1994).

[9] J. J. Huang, H. P. Pan, Y. C. Wan, "An algorithm for cooperative learning of Bayesian network structure from data," Lecture Notes in Computer Science. 3168, 86-94(2005).

[10] A. Onisko, M. J. Druzdzel, H. Wasyluk, "Learning Bayesian network parameters from small data sets: Application of Noisy-OR gates," International Journal of Approximate Reasoning. 27(2), 165-182(2001).

[11] A. Feelders, C. L. Van, "Learning Bayesian network parameters under order constraints," International Journal of Approximate Reasoning. 42 (1-2), 37-53(2006).

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[12] A. Porwal, E. J. Carranza, M. Hale, "Bayesian network classifiers for mineral potential mapping," Computers and Geosciences. 32 (1), 1-16(2006).

[13] Thomas Liebig, Christine Körner, Michael May, "Fast Visual Trajectory Analysis Using Spatial Bayesian Networks", ICDMW, 2009, 2013 IEEE 13th International Conference on Data Mining Workshops, 2013 IEEE 13th International Conference on Data Mining Workshops 2009, pp. 668-673, doi:10.1109/ICDMW.2009.44

[14] Walker, Arron R., Pham, Binh L., & Moody, Miles P. (2005) Spatial Bayesian learning algorithms for geographic information retrieval. In 13th annual ACM international workshop on Geographic information systems, November 2005, Bremen, Germany.

[15] Li L1, Wang J, Leung H, Zhao S. A Bayesian method to mine spatial data sets to evaluate the vulnerability of human beings to catastrophic risk. Risk Anal. 2012 Jun;32(6):1072-92. doi: 10.1111/j.1539-6924.2012.01790.x. Epub 2012 Mar 9.

[16] R. Kimball, "The Data Warehouse Toolkit" John Wiley and Sons, New York, 1996.

[17] W.H.Inmon "Building the Data Warehouse John Wiley",1996.

[18] Bedard, T. Merrett, and J. Han, "Fundamentals of Geospatial Data Warehousing for Geographic Knowledge Discovery", H. Miller and J. Han (eds.), In Geographic Data Mining and Knowledge Discovery, Taylor and Francis, 2001.

[19] Han, J., Stefanovic, and K. Koperski, "Selective Materialization: An Efficient Method for Spatial Data Cube Construction", Proceedings Pacific-Asia Conference in Knowledge Discovery and Data Mining, PAKDD'98, Melbourne, Australia. [VAL 87]

[20] S. Chawla, S. Shekhar, W. Wu, and U. Ozesmi, "Modeling spatial dependencies for mining geospatial data: An introduction," In Geographic data mining and Knowledge Discovery(GKD), H. Miller and J. Han, Eds. Taylor and Francis, 2001.

[21] VALDURIEZ P., "Join indices", ACM Trans. on Database Systems, 12(2); 218-246, June 1987.

[22] Zeitouni K.,Yeh L., Aufaure M-A, Join indices as a tool for spatial data mining, Int. Workshop on Temporal, Spatial and Spatio-Temporal Data Mining, Lecture Notes in Artificial Intelligence n° 2007, Springer, pp 102-114, Lyon, France, September 12-16, 2000.

[23] Elzbieta Malinowski, Esteban Zimányi Spatial Hierarchies and Topological Relationships in the Spatial MultiDimER Model Database: Enterprise, Skills and Innovation Volume 3567 of the series Lecture Notes in Computer Science pp 17-28 (2005)

[24] W. H. Inmon. Building the Data Warehouse. QED Technical Publishing Group, Wellesley, Massachusetts, 1992.

[25] Pumain D., Saint-Julien Th., 1997 (réédité en 2004) : L'analyse spatiale. 1: Localisations dans l'espace. Armand Colin, Coll. « Cursus », 167 p.