

Provided by the author(s) and University College Dublin Library in accordance with publisher policies. Please cite the published version when available.

Title	Mining Spatio-temporal Data at Different Levels of Detail
Author(s)	Camossi, Elena; Bertolotto, Michela; Kechadi, Tahar
Publication Date	2008-05-08
Publication information	Bernard, L., Friis-Christen, A., and Pundt, H. (eds.) The European Information Society : taking geo-information science one step further
Publisher	Springer-Verlag
Link to publisher's version	http://dx.doi.org/10.1007/978-3-540-78946-8_12
This item's record/more information	http://hdl.handle.net/10197/1442
DOI	http://dx.doi.org/10.1007/978-3-540-78946-8_12

Downloaded 2013-09-13T15:58:10Z

Some rights reserved. For more information, please see the item record link above.



Mining Spatio-temporal Data at Different Levels of Detail

Elena Camossi, Michela Bertolotto, Tahar Kechadi

School of Computer Science & Informatics, University College Dublin, Belfield, Dublin 4, Ireland. {elena.camossi, michela.bertolotto, tahar.kechadi}@ucd.ie

Abstract

In this paper we propose a methodology for mining very large spatio-temporal datasets. We propose a two-pass strategy for mining and manipulating spatio-temporal datasets at different levels of detail (i.e., granularities). The approach takes advantage of the multi-granular capability of the underlying spatio-temporal model to reduce the amount of data that can be accessed initially. The approach is implemented and applied to real-world spatio-temporal datasets. We show that the technique can deal easily with very large datasets without losing the accuracy of the extracted patterns, as demonstrated in the experimental results.

1 Introduction

Recently, it has been estimated that 80% of the available datasets have spatial components (Fayyad and Grinstein, 2001), and are often related to some temporal aspects. Such a considerable amount of information needs suitable analysis techniques to be applied correctly. In the last few years, several systems providing an integrated approach for the management of spatial and temporal information have been proposed (e.g., Chen and Zaniolo, 2000, Güting et al., 2000, Huang and Claramunt, 2002).

The application of knowledge management tailored to the exploitation of implicit semantics of spatio-temporal data has emerged as the key technology to address the application of spatio-temporal data mining techniques and algorithms to real-world problems. Spatio-temporal data mining is a user-centric, interactive process, where data mining experts and domain experts work closely together to gain insight on a given problem. Several open issues have been identified ranging from the definition of the mining techniques capable of dealing with spatio-temporal

poral information to the development of effective methods for interpreting and presenting the final results.

In this study, we focus on a specific data mining technique that deals with clustering. Clustering is one of the fundamental techniques in data mining. It groups data objects into clusters based on some similarity or distance measures. These clusters contain information found in the data that describes similar objects and their relationships. The goal is to optimise similarity within a group of objects and dissimilarity between the groups in order to identify interesting structures in the underlying data. While the complexity of spatio-temporal clustering is far higher than its traditional counterpart, the ideas behind it are similar i.e., it focuses either on characteristic features of objects in a spatio-temporal region or on the spatio-temporal characteristics of a set of objects (Ng and Han, 1994).

The mining process for spatio-temporal data is complex in terms of both the mining efficiency and the complexity of patterns that can be extracted from spatio-temporal datasets (Roddick and Lees, 2001). The reason is that the attributes of the neighbouring patterns (i.e., close in either space or time or both) may have significant influence on a pattern and should also be considered. Therefore, new techniques are required to efficiently and effectively mine these datasets. The main problem for analysing spatio-temporal data is the size of the data. Today's GIS systems are collecting Gigabytes and even Terabytes of data each day. So the major goal for such a strategy is to process these datasets within a reasonable response time and memory space, without affecting the accuracy of the findings.

In this paper we propose a spatio-temporal clustering technique to deal with the data at different levels of detail, i.e., granularities, to improve the algorithm efficiency. Such a technique relies on a hierarchical multi-granular model in which datasets are generalised to generate less detailed representations of reduced size. Thus, the mining can first be applied to the reduced dataset, and then refined only for those objects, which have been filtered through the first step. In other words, the mining can be further deepened on spatial areas or temporal intervals of interest. The corresponding objects are converted at finer spatial and temporal granularity before applying the mining. Our approach handles the data at different levels of detail both from a spatial and a temporal point of view. The conversions of data at different levels of detail are performed by applying the operators available in the underlying multi-granular spatio-temporal model, whose definitions are described in (Camossi et al., 2006).

The paper is organised as follows. In Section 2 we present recent related work. In Section 3 we describe a multi-granular spatio-temporal model that enables the conversion of spatio-temporal data at different levels of detail. In Section 4 we introduce the spatio-temporal data mining system and show how we apply it to spatio-temporal data represented at higher spatial and temporal levels of granularity. In Section 5 we discuss some experimental results. Finally, Section 6 concludes the paper and outlines future research directions.

2 Related Work

The proposals for the integrated management of spatio-temporal information can be mainly classified into: temporal extensions of GIS (Claramunt and Thériault, 1995, Langran, 1992); extensions of relational, object relational (Chen and Zaniolo, 2000) and object oriented standards (Griffiths et al., 2004, Huang and Claramunt, 2002); algebraic frameworks for moving points and regions (Güting et al.,

2000); and independent frameworks (Tryfona and Jensen, 1999, Worboys, 1994).

Recent systems have addressed the issues related to multi-granularity, multi-resolution and multiple representations of spatial (Balley et al., 2004, Fonseca et al., 2002, Kulik et al., 2005, Vangenot, 2001) and spatio-temporal data (Bittner, 2002, Camossi et al., 2006, Claramunt and Jiang, 2000, Hornsby and Egenhofer, 2002), Hurtado and Mendelzon, 2001, Khatri et al., 2002). In particular, Claramunt and Jiang (2000) defined nested hierarchies for modelling space and time from which quantitative information about spatio-temporal relationships are obtained. Khatri et al. (2002) extended a semantic formalism to support the specifications of spatio-temporal data at multiple granularities, relying on the concepts of temporal indeterminacy and spatial imprecision. The resulting model and the granularity systems described are effective for data specification. In (Camossi et al., 2006) a framework enabling the conversion of spatio-temporal values at different spatial and temporal granularities is defined as extension of the ODMG data model (Cattel et al., 1999). In the spatial domain, Fonseca et al. (2002) and Kulik et al. (2005) proposed the use of anthologies to multi-resolution.

The progressive application of data mining techniques for spatio-temporal data to improve efficiency is discussed in (Mennis and Liu, 2005, Tsoukatos and Gunopulos, 2001). Tsoukatos and Gunopulos (2001) presented an incremental algorithm for discovering frequent spatio-temporal sequences by decomposing the search space in a hierarchical structure, addressing its application to multi-granular spatial data. Mennis and Liu (2005) discussed multi-level association rule mining of spatio-temporal data, i.e., mining of rules at varying levels of a concept hierarchy to fit the best resolution for the rule. Hierarchical data mining is discussed also for spatial (Koperski, 1999, Shahabi et al., 2001) and temporal (Abraham and Roddick, 1999) data separately. Recently, there has been a growing interest in the application of wavelet transforms in some processes of data mining (Li et al., 2002, Shahabi et al., 2001).

3 Multi-granular representation of spatio-temporal data

In this section we describe the data model for the representation of data at multiple spatio-temporal granularities used in our mining approach. The model relies on the work presented in (Camossi et al., 2006), where the ODMG type system (Cattel et al., 1999) has been extended to enable the representation and the conversion of spatio-temporal object attributes at different levels of details, for both the spatial and the temporal dimensions. The same set of conversions has been applied in the definition of an object-relational spatio-temporal multigranular model (Bertino et al., 2005). In this paper we follow the object-relational approach, instead of the full object oriented approach, because it is adopted by current commercial DBMS. Furthermore, like most of them (e.g., ORACLE™, 2008, PostgreSQL, 2008), the model applies an integrated approach for the representation of geometric aspects of data. In the following, we first present the notion of spatial and temporal granularities supported by the model; then, we describe how multi-granular spatio-temporal data can be represented and converted.

3.1 A spatio-temporal multi-granular data model

The data model supports the definition of temporal granularity formalised by Bet-

tini et al. (2000), which is commonly adopted by the temporal databases and reasoning community, and integrates the notion of spatial granularity compliant with the formalization of *stratified map spaces* proposed by Stell and Worboys (1998). Temporal and spatial granularities are specified as mappings from an index set to the power set of the *TIME* and the *SPACE* domains, respectively. *TIME* is totally ordered. The supported *SPACE* domain is 2-dimensional (i.e., a proper subset of R^2). For instance, *days*, *weeks*, *years* are temporal granularities; *meters*, *kilometres*, *feet*, *yards*, *provinces* and *countries* are spatial granularities. Each portion of the temporal and spatial domain corresponding to a granularity mapping is referred to as a (temporal or spatial) *granule*. Spatial granularities can include 2-dimensional granules (e.g., units of area: *m²*, *acre*, etc.; administrative boundaries classifications: *municipalities*, *countries*, etc.), or in 1-dimensional granules (e.g., measures of length: *km*, *mile*, etc.; map scales: 1:24 000, 1:62 500, etc.). Granules give the validity bounds of spatio-temporal for the definition of spatio-temporal values. For instance, we can say that a value reporting the measure of the daily temperature in Dublin is defined for the first and the second of January 2000, and so on. "01/01/2000", "02/01/2000", and "Dublin" are textual labels that univocally identify two temporal and one spatial granule. Granules of the same granularity cannot overlap. Moreover, non-empty temporal granules must preserve the order given by the index set.

Spatial and temporal granularities are related by the *finer-than* relationship. Such a relationship formalises the intuitive idea that different granularities correspond to different partitions of the domain, and that, given a granule of a granularity *G*, usually a granule of a coarser granularity exists that properly includes it. For example, granularity *days* is finer than *months*, and granularity *months* is finer than *years*. Likewise, *municipalities* is finer than *countries*. If a granularity *G* is finer than *H*, we also say that *H* is *coarser-than* *G*. According to the finer-than relationship, spatial and temporal granularities are related to form two directed graphs, usually two lattices.

Beyond the conventional relational and object-relational database values, the database schemas can include spatial, temporal, and spatio-temporal values. 2-dimensional geometric vector features (i.e., points, lines, and polygon) can then be represented. Multi-granular spatial and temporal data are uniformly defined by instances of two parametric types, *spatial* and *temporal*, which are specified according to granularities (spatial and temporal, respectively) and an inner conventional (i.e., without spatio-temporal characteristics) or geometric type.

The model enables the conversion of multigranular spatio-temporal data at different granularities, to improve or reduce the level of detail employed for data representation. Granularity conversions are crucial in order to represent data at the most appropriate level of detail for a specific task, and enable consistent comparisons of data defined in the schema at different granularities, improving the expressive power of spatio-temporal query languages. Granularity conversions enable to apply different conversion semantics.

The conversion of multi-granular geometrical features is obtained through compositions of model-oriented and cartographic map generalisation operators (Muller et al., 1995) that guarantee topological consistency (Bertolotto, 1998, Saalfeld, 1999), an essential property for data usability, and refinement operators that perform the inverse functions. Such operators can be classified with respect to the semantics of the conversion performed: *contraction* and *thinning* operators reduce the dimension of vector features, whereas *expansion* operators increase their dimension; *merge* operators merge adjacent features of the same dimension into a single one, while *splitting* operators subdivide single features in adjacent features of the same dimension; *abstraction* and *simplification* operators discard isolated

features from polygons and remove shape points from a line, respectively, whereas *addition* operators add isolated features to polygons and shape points to lines.

On the other hand, to retrieve for instance the annual trend of a phenomenon with a daily frequency (e.g., the national values of sales in shops located in several countries, the model supports also the conversion of quantitative (i.e., not geometrical) attribute values supported for both temporal and spatial data. These conversions are classified in families according to the semantics of the operation performed (Camossi et al., 2006): *selection* (e.g., projection); *aggregation* (e.g., sum, average); *restriction*, by which, if a granular value assumes value v in a granule g , value v also refers to any finer granule g' included in g ; *splitting*, which subdivides each coarser value among the finer granules included in it either uniformly (i.e., all the finer values are the same), or according to non-uniform distribution.

The given set of granularity conversions can be extended with user-defined granularity conversions that are specified as class methods in a database schema. Granularity conversions have been proved to return legal values of the type system defined, and to preserve the semantics of the spatio-temporal data represented (Camossi et al., 2006).

3.2 Multi-granularity to Improve Mining

In this paper, we take advantage of the multi-granularity support provided by the data model to enhance the effectiveness of the clustering algorithm. In particular, the mining process can benefit of multi-granularity in different ways.

First of all, multi-granularity enables to apply the mining to data represented at different levels of detail, e.g., semantically homogeneous data coming from different sources. In this case, data can be converted into uniform spatial and temporal granularities before applying the mining process. The level of detail is chosen in order to represent the specific dataset. Usually the choice falls on the *greatest lower bound* (*glb*), or the *least upper bound* (*lub*), of the spatial and temporal granularity available. Given two granularities G and H of the same type (i.e., either spatial or temporal), $glb(G,H)$ is the coarsest granularity K (not necessarily different from G and H) among the granularities finer than both G and H . By contrast, $lub(G,H)$ is the finest granularity J (not necessarily different from G and H) among the granularities coarser than both G and H .

Then, once the level of detail used for the representation is uniform for the whole dataset, granularity conversions are applied before the refinement process. Indeed, spatio-temporal data are pre-processed for reducing the size of the starting dataset, i.e., data are converted to coarser spatial and temporal granularities. This conversion allows us to focus on the relevant dataset, which is, in general, much smaller than the original data, hence, improving response time of the overall mining process. The choice of the level of detail can be iterative, and depends on a trade-off between mining efficiency and maximum detail required by the mining process. Finally, the conversion depends on the generalisation process used for a given dataset. Once the semantics for generalising certain dimensions or attributes of the data has been defined, the conversion is straightforward and mainly for the model defined above. Therefore, the mining process needs only the level of accuracy as an input parameter and the conversion and even the number of levels of detail that need to be explored is done automatically through hierarchical navigation. The way that these levels are explored depends on the type of the algorithm implemented.

After the application of the clustering algorithm, the selected spatio-temporal

data of interest are converted into finer granularities, for more detailed representation, once a deeper analysis is required on significant data. In the first case, granularity conversions are applied globally, to the whole dataset to materialise those objects, which will be accessed frequently and in more detail. In the second case, granularity conversions are applied locally, zooming in on specified and restricted pieces of information, whenever the user asks for a more detailed mining of such data, specified with respect to a given spatio-temporal area. In both cases, spatio-temporal data are converted to different granularities without losing information. Indeed, the conversions are performed by applying the granularity conversions supported by the data model that preserve semantics and then usability of the data.

4 Proposed System

To address the issues of mining and managing spatio-temporal datasets we have proposed a 2-layer system architecture (Bertolotto et al., 2007, Compieta et al., 2007) including a mining layer and a visualisation layer. The mining layer implements a mining process along with the data preparation and interpretation steps. For instance, the data may need some cleaning and transformation according to possible constraints imposed by some tools, algorithms, or users. The interpretation step consists of visualising the selected models returned during the mining phase to effectively study the application behaviour. The interpretation is carried out in the visualisation layer. More details on the visualisation tools can be found in (Bertolotto et al., 2007, Compieta et al., 2007). In the next section we will focus on the mining strategy implemented in the mining layer.

4.1 2-Pass Strategy

To reduce the amount of memory and computational complexity that these data spaces require without affecting the information presented by the data, the first task in our strategy is to find the data points that are most similar according to their static (non spatial and temporal) attributes. This part of the strategy is the key to the whole success of the generalisation process, so that we do not lose any important information that might have an adverse effect on the results. To further reduce the complexity in space of the algorithm, the raw datasets are pre-processed in order to obtain, through granularity conversions defined in Section 3, a coarser representation of their spatio-temporal dimensions. Since the granularity conversion preserves the semantics of data (Camossi et al., 2006), the application of spatio-temporal mining algorithms to coarser representation does not affect the algorithm outputs.

The second task is to cluster these groups of closely related data points in a meaningful way to produce a new (meta-)dataset suitable and acceptable for further analysis (i.e., models, patterns, rules, etc.).

4.2 First Pass

The algorithm for this first pass produces clusters of data points that are closely related. The goal here is to produce new data objects, where each object represents

one cluster of raw data. Therefore, the main objective is to reduce the size of the initial data without losing any relevant information. Figure 1 shows a high level view of the steps carried out by the algorithm for this first phase of this approach. It is important to note that only the data points that have a very high similarity between them will be grouped together. As a result, the new dataset is much smaller than the original data. It contains more information about individual clusters. This will help the clustering performed during the second pass.

This pass is basically implementing the generalisation and conversion model defined above. The process of exploring the generalised data and its conversion either from top-to-bottom or bottom-to-top is linear. Usually the generalisation process is implemented as a tree structure, which is efficient in exploring relevant branches and the memory space needed to store them. In this phase, we access the higher-level generalised data. The second pass will deal with the detail when necessary.

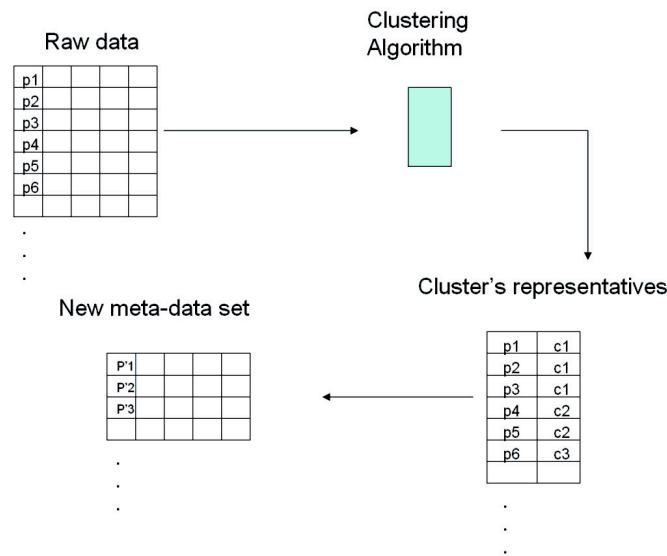


Figure 1: Step-by-step view of the first pass of the strategy.

4.3 Second Pass

The second pass involves clustering the tightly grouped data points from the first pass to produce a new representation of the data (meta-dataset). This meta-dataset should be reduced by a certain degree of magnitude so that it can now be analysed and mined more easily. In Figure 2, a 2-D example shows how a larger dataset is processed to create a much smaller meta-dataset. There are locations in the space that are highly similar; these are represented within each of the small location groups (small circular shapes). It is important that no location group overlaps with another so that the integrity of the data is not affected. The next step shows data

mining on the meta-data using clustering for an example.

The clustering technique proposed for this second phase of the strategy will be DBSCAN (Ester et al., 1996). It is a density-based clustering algorithm that produces disjoint clusters, in which the number of clusters is automatically determined by the algorithm. It is relatively resistant to noise (as it detects noisy data and outliers) and can handle clusters of arbitrary shapes and sizes. The main reason for choosing DBSCAN is twofold: 1) to illustrate our methodology and our conversion model, and 2) while DBSCAN is not highly scalable; it is interesting to study its performance on very large datasets using our methodology as from our first phase the amount of tiny clusters representing highly similar data is very large, and we would like to take advantage of finding the regions that are very similar. These regions can then form clusters that will present a new compact representation of the dataset.

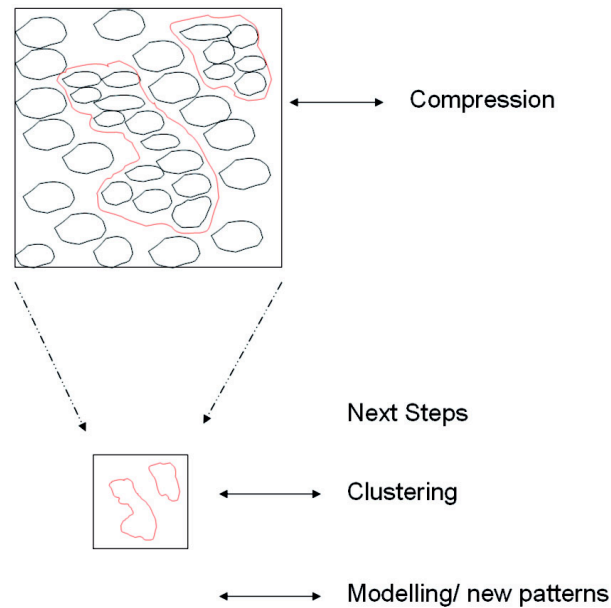


Figure 2: A 2-d example of dataset compression.

The next step is to mine this new representation of the dataset. The space and computational complexities for these algorithms have been reduced greatly from the original data. This strategy and mainly the mining algorithm is also suitable for interactive data mining and visualisation since it is so quick and efficient that it can be incorporated in a visualisation tool of the data. The data can be explored and analysed using this approach interactively and with ease as shown in the next section.

5 Preliminary Experimental Results

We have implemented a 2-pass strategy that uses the DBSCAN algorithm for the

mining process. The technique has been implemented within a data-mining engine and includes also a visualisation layer for interactive data interpretation. The experiments conducted so far were obtained from the Hurricane Isabel dataset (National Hurricane Center, 2003), which is a proper instance for geographical spatio-temporal dataset. Figure 3 lists different variables contained in the dataset and for more details about these variables we refer the reader to (National Hurricane Center, 2003).

Variable Descriptions

Variable	Description	Min/Max	Units
QCLOUD	Cloud Water	0.00000/0.00332	kg/kg
QGRAUP	Graupel	0.00000/0.01638	kg/kg
QICE	Cloud Ice	0.00000/0.00099	kg/kg
QRAIN	Rain	0.00000/0.01132	kg/kg
QSNOW	Snow	0.00000/0.00135	kg/kg
QVAPOR	Water Vapor	0.00000/0.02368	kg/kg
CLOUD	Total cloud (QICE+QCLOUD)	0.00000/0.00332	kg/kg
PRECIP	Total Precipitation (QGRAUP+QRAIN+QSNOW)	0.00000/0.01672	kg/kg
P	Pressure; weight of the atmosphere above a grid point	-5471.85791/3225.42578	Pascals
TC	Temperature	-83.00402/31.51576	Degrees Celsius
U	X wind component; west-east wind component in model coordinate, positive means winds blow from west to east	-79.47297/85.17703	
V	Y wind component; south-north wind component in model coordinate, positive means winds blow from south to north	-76.03391/82.95293	
W	Z wind component; vertical wind component in model coordinate, positive means upward motion	-9.06026/28.61434	

Figure 3: Description of the dataset layers.

All variables are real-valued (4 bytes) and were observed along 48 time steps (hourly-sampled), in a space having $500 \times 500 \times 100 = 25 \times 10^6$ total points. So, each variable in each time step is stored in a different file, resulting in 624 files of 100MB each. This raw data can be represented by the following parameters; the number of time steps (Nts), the number of data points (N), and the number of static parameters (Nsp). Nts = 48 time steps, N = 25×10^6 data points, and Nsp = 13. This fine fragmentation allows for great flexibility in choosing different subset of data for each mining task.

Figure 4 shows one of the clusters we extracted, whose shape resembles the shape of the hurricane or one of its features. In Figure 4, DBSCAN algorithm outputs a spherical type cluster that represents the shape of the Hurricane's eye for different values of pressure. The eye is clearly visible in the low, where the cluster represents high values for pressure in pink with the hole in the middle representing very low values for pressure. These clusters provide some clues about direction or strength of the hurricane. We can track and represent in real time the movement of the hurricane eye over time by clustering different time steps of the dataset.

The application of the mining algorithm on the reduced dataset produces results visually comparable to those obtained with the fully detailed one, and with improved efficiency in response time and memory occupation. The results we ob-

tained on the test-bed application (i.e., the Hurricane Isabel) are very promising, and this technique suites very well interactive environments. Our technique is designed according to the multi-level granularity model explained above. In this paper, we presented our results by using a 2-pass (i.e., 2-level) algorithm. The first level generalises the dataset to reduce its size and complexity. The second pass refines only the data of interest. However, these improved results may depend on the specific dataset. For instance, if the final patterns cover different objects, which were not identified to be neighbours in the first pass, the cost will be higher as one has to explore different spatio-temporal regions to refine the final results. This can be solved by adapting our algorithm to support multiple levels of clustering relying on the model defined above. We are currently implementing a version for multi-level clustering using decision-tree approach.

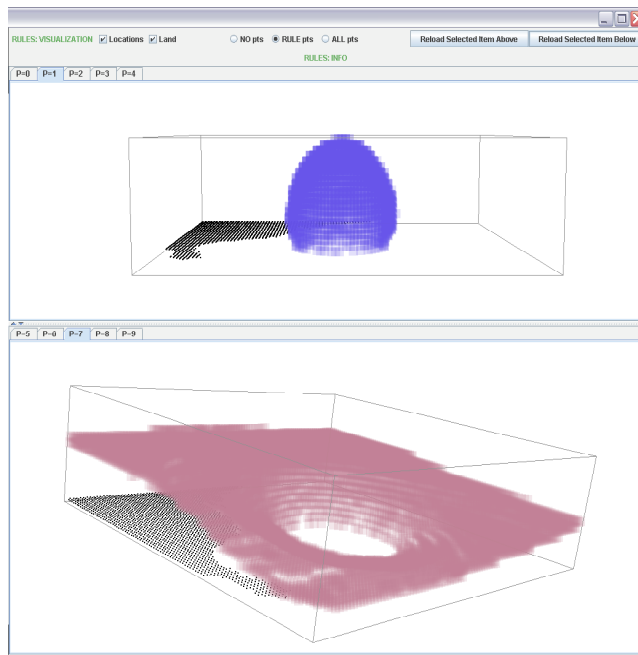


Figure 4: The eye of the Hurricane in isolation. This is represented by one cluster.

6 Conclusion

The approach proposed in this paper is different from the approaches presented in the literature (Abraham and Roddick, 1999, Koperski, 1999, Mennis and Liu, 2005, Tsoukatos and Gunopulos, 2001) with respect to the specific data mining problem addressed, and mainly the use of multi-granularity concept to both be able to design scalable technique for data mining and analysis and speed up the process of the mining and its accuracy. The work presented in (Tsoukatos and

Gunopulos, 2001) focuses on mining frequent patterns, while (Abraham and Roddick, 1999, Koperski, 1999, Mennis and Liu, 2005) address the mining of association rules, meta-rules and classification. (Tsoukatos and Gunopulos, 2001) uses spatial granularities defined according to boundary regions, and the operator supported to perform granularity conversions is region merge. Likewise, only spatial concept hierarchy are supported in (Abraham and Roddick, 1999, Koperski, 1999, Mennis and Liu, 2005). In all these projects, spatial granularities are employed for rules representations. Instead, we focus on clustering, and the multi-granularity concept is used to reduce the size of the datasets, mainly at the beginning. Furthermore, we apply multi-granularity for both the spatial and the temporal domains, supporting a wide range of granularity conversions, specifically designed to preserve data usability. We will extend this approach to other clustering techniques and also we will study their effectiveness in real-world environments. Moreover, we have planned further experimentations considering different spatio-temporal datasets to the test of the efficiency of the approach.

Acknowledgement

The work of Elena Camossi is supported by the Irish Research Council for Science, Engineering and Technology.

References

- Abraham T., Roddick J.F. (1999) Incremental Meta-Mining from Large Temporal Datasets. *Advances in Database Technologies*, In Proc. of the *1st Int'l Workshop on Data Warehousing and Data Mining*, Springer-Verlag Berlin. LNCS 1552:41-54.
- Balley S., Parent C., Spaccapietra S. (2004) Modelling Geographic Data with Multiple Representations. *International Journal of Geographical Information Science*, Taylor & Francis. 18(4):327-352.
- Bertino E., Cuadra D., Martínez P. (2005) An Object-Relational Approach to the Representation of Multi-granular Spatio-Temporal Data. In Proc. of the *17th Int'l Conf. on Advanced Information Systems Engineering*, Springer-Verlag Berlin. LNCS 3520:119-134.
- Bertolotto M. (1998) Geometric Modeling of Spatial Entities at Multiple Levels of Resolution. Ph.D. Thesis, Università degli Studi di Genova, Italy.
- Bertolotto M., Di Martino S., Ferrucci F., Kechadi T. (2007) A Visualisation System for Collaborative Spatio-Temporal Data Mining. *International Journal of Geographical Information Science*, Taylor & Francis. 21(7): 895-906.
- Bettini C., Jajodia S., Wang X. (2000) Time Granularities in Databases, Data Mining, and Temporal Reasoning, Springer-Verlag Berlin.
- Bittner T. (2002) Reasoning about qualitative spatio-temporal relations at multiple levels of granularity. In Proc. of the *15th European Conf. on Artificial Intelligence*, IOS Press. 317-321.
- Camossi E., Bertolotto M., Bertino E. (2006) A multigranular Object-oriented Framework Supporting Spatio-temporal Granularity Conversions. *International Journal of Geographical Information Science*. Taylor & Francis. 20(5): 511-534.

- Cattel R., Barry D., Berler M., Eastman J., Jordan D., Russel C., Schadow O., Stanienda T., Velez F (1999). The Object Database Standard: ODMG 3.0. Morgan-Kaufmann.
- Claramunt C., Thériault M. (1995) Managing Time in GIS: an event oriented approach. In Proc. of the *Int'l Workshop on Temporal Databases: Recent Advances in Temporal Databases*, Springer-Verlag. 23-42.
- Claramunt C., Jiang B. (2000) Hierarchical Reasoning in Time and Space. In Proc. of the *9th Int'l Symposium on Spatial Data Handling*. 41-51.
- Compieta P., Di Martino S., Bertolotto M., Ferrucci F., Kechadi T. (2007) Exploratory spatio-temporal data mining and visualization. *Journal of Visual Languages and Computing*, Elsevier. 18(3):255-279.
- Chen C.X., Zaniolo C. (2000) SQL: A Spatio-Temporal Data Model and Query Language. In Proc. of *19th Int'l Conf. on Conceptual Modeling / the Entity Relational Approach*. Springer-Verlag Berlin. LNCS 1920:96-111.
- Ester M., Kriegel H.-P., Sander J., Xu X. (1996) A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. In Proc. of the *2nd Int'l Conf. on Knowledge Discovery and Data Mining*. 226-231.
- Fayyad U.M., Grinstein G.G. (2001) Introduction. *Information Visualization in Data Mining and Knowledge Discovery*, Los Altos, CA: Morgan Kaufmann. 1-17.
- Fonseca F., Egenhofer M.J, Davis C., Cãmara G. (2002) Semantic Granularity in Ontology Driven Geographic Information Systems. *Annals of Mathematics and Artificial Intelligence, Special Issue on Spatial and Temporal Granularity*. 36(1-2).
- Griffiths T., Fernandes A.A.A., Paton N.W., Barr R. (2004). The Tripod spatio-historical data model. *Data Knowledge and Engineering*, Elsevier. 49(1): 23-65.
- Gütting R.H., Bhölen M.H., Erwig M., Jensen C.S., Lorentzos N.A., Shneider M., Vazirgiannis M. (2000) A Foundation for Representing and Querying Moving Objects. *ACM Transaction On Database Systems*, 25:1-42.
- Hornsby K., Egenhofer M.J. (2002) Modeling Moving Objects over Multiple Granularities. *Annals of Mathematics and Artificial Intelligence*. Special Issue on Spatial and Temporal Granularity. Kluwer Academic Press. 36(1-2):177-194.
- Hurtado C.A., Mendelzon A.O. (2001) Reasoning about summarizability in Heterogeneous Multidimensional Schemas. In Proc. of the *8th Int'l Conf. on Database Theory*. 375-389.
- Huang B., Claramunt C. (2002) STOQL: An ODMG-based Spatio-Temporal Object Model and Query Language. In Proc. of the *10th Int'l Symposium on Spatial Data Handling*, Springer-Verlag Berlin. 225-237.
- Khatri V., Ram S., Snodgrass R.T., O'Brien G. (2002) Supporting User Defined Granularities and Indeterminacy in a Spatio-temporal Conceptual Model. *Annals of Mathematics and Artificial Intelligence*. Special Issue on Spatial and Temporal Granularity, 36(1):195-232.
- Koperski K.(1999) A Progressive Refinement Approach to Spatial Data Mining. Ph.D. Thesis, Simon Fraser University, Canada.
- Kulik L., Duckham M., Egenhofer M.J. (2005) Ontology driven Map Generalization. *Journal of Visual Language and Computing*, 16(3):245-267.
- Langran G. (1992) Time in Geographic Information Systems. Taylor & Francis.
- Li T., Li Q., Zhu S., Ogihara M. (2002) A Survey on Wavelet Applications in Data Mining. *ACM SIGKDD Explorations Newsletter*. 4(2):49-68.
- Mennis J., Liu J.W. (2005) Mining Association Rules in Spatio-Temporal Data: An Analysis of Urban Socioeconomic and Land Cover Change. *Transactions in*

- GIS, Blackwell Publishing. 9(1):5–17.
- Muller J.-C., Lagrange J.P., Weibel R. (eds.) (1995) GIS and Generalization: methodology and practice. Taylor and Francis.
- National Hurricane Center (2003), *Tropical Cyclone Report: Hurricane Isabel*, <http://www.tpc.ncep.noaa.gov/2003isabel.shtml>.
- Ng R.T., Han J. (1994) Efficient and Effective Clustering Methods for Spatial Data Mining. In Proc. of the 20th Int'l Conf. on Very Large Data Bases. 144-155.
- ORACLE™ (2008), Oracle Corp. <http://www.oracle.com>. Last date accessed: 01/2008.
- PostgreSQL (2008), PostgreSQL Inc. <http://www.postgresql.org>. Last date accessed: 01/2008.
- Roddick J.F., Lees B.G. (2001) Paradigms for Spatial and Spatio-Temporal Data Mining. *Geographic Data Mining and Knowledge Discovery*. Taylor and Francis. 33-50.
- Saalfeld A. (1999) Topologically consistent line simplification with the Douglas-Peucker algorithm. *Cartography and Geographic Information Science*. 26(1):7-18.
- Shahabi C., Chung S., Safar M., Hajj G. (2001) 2D TSA-tree: A Wavelet-Based Approach to Improve the Efficiency of MultiLevel Spatial Data Mining. In Proc. of the 13th Int'l Conf. on Scientific and Statistical Database Management. 59-68.
- Stell J.G., Worboys M. (1998) Stratified Map Spaces: A Formal Basis for Multi-Resolution Spatial Databases. In Proc. of the 8th Int'l Symposium on Spatial Data Handling. 180-189.
- Tsoukatos I., Gunopulos D. (2001) Efficient Mining of Spatiotemporal Patterns. In Proc. of the 7th Int'l Symposium on Spatial and Temporal Databases. LNCS 2121:425-442.
- Tryfona N., Jensen C.S. (1999) Conceptual Modeling for Spatiotemporal Applications. *Geoinformatica*, Springer Netherlands. 3(3):245-268.
- Vangenot C. (2001) Supporting Decision-Making with Alternative Data Representations. *Journal of Geographic Information and Decision Analysis*. 5(2):66-82.
- Worboys M. (1994) A Unified Model for Spatial and Temporal Information. *The Computer Journal*, Oxford University Press. 37(1):26-34.