

GeoSTAT – A system for visualization, analysis and clustering of distributed spatiotemporal data

Maxwell Guimarães de Oliveira, Cláudio de Souza Baptista

Laboratory of Information Systems – Computer Science Department
Federal University of Campina Grande (UFCG)
Av. Aprígio Veloso 882, Bloco CN, Bairro Universitário – 58.429-140
Campina Grande – PB – Brazil

maxmcz@gmail.com, baptista@dsc.ufcg.edu.br

***Abstract.** Nowadays, there is a considerable amount of spatiotemporal data available on the web. The visualization of these data requires several visual resources which helps users to have a correct interpretation of the data set. Furthermore, the use of data mining algorithms has proven relevant in helping the exploratory analysis of spatiotemporal data. This paper proposes the GeoSTAT (GEOgraphic SpatioTemporal Analysis Tool), a system that includes spatial and temporal visualization techniques and offers a spatiotemporal adaptation of clustering algorithms provided by the Weka data mining toolkit. A case study was realized to demonstrate the end-user experience and some advantages achieved using the proposed system.*

1. Introduction

Nowadays, there is a considerable volume of spatiotemporal data available in a variety of media types, especially on the Internet. Among so much information, it is necessary to provide decision support systems and analytics, which can help decision making users to extract relevant knowledge, intuitively and quickly, such as the prediction of future events, for instance.

Visualization techniques are widely known as being powerful in the decision making domain [Johnston 2001], since they take advantage of human capabilities to rapidly notice and interpret visual patterns [Andrienko et al. 2003][Kopanakis and Theodoulidis 2003]. However, we know that the spatial visualization resources supplied by most of the existing geographic information systems are not enough for decision support systems [Bédard et al. 2001].

The visualization of spatiotemporal data is a complex task that requires the use of appropriate visual resources that allow users to have a correct interpretation of the information under analysis. Visualization and analysis of spatiotemporal data are tasks that have been gaining prominence in several areas, such as biology, electrical power transmission, urban traffic, criminology, and civil construction. This cross domain utilization is especially due to the widespread use of devices that capture the geographic location, generating large amounts of information concerning the time and space, such as the trajectory of mobile objects, fire spots, dengue spots, atmospheric discharges, and criminality maps.

According to Andrienko et al. [Andrienko et al. 2010b], it is necessary to deal with the time in an efficient manner, when performing spatiotemporal visualization. The understanding that space and time are inseparable and that there is nothing spatial that is

not temporal must permeate the research in spatiotemporal visualization. A reasonable solution in visualization and analysis of spatiotemporal data should offer, at least: resources for treating both the spatial and temporal dimensions (spatiality and temporality); domain independence (generality), freedom for the user to handle the visualized data and apply filters (flexibility); connection with several data sources in a practical and efficient manner (interoperability); and data mining based on spatiotemporal clustering (mining).

It is essential to provide to the users resources to handle both the spatial and the temporal dimensions in a spatiotemporal data analysis system. The singularities in any of these dimensions must not be discarded because they may reveal implicit relationships which match the reality of the analyzed data.

Furthermore, the use of spatiotemporal data mining algorithms, integrated with modern data visualization techniques, improves the usability for the decision maker when analyzing large spatiotemporal datasets.

Nonetheless, the majority of existing spatiotemporal visualization systems do not address appropriately the temporal dimension, as they focus only the spatial visualization. Therefore, an important research issue is how to offer temporal manipulation resources that, used with the spatial data manipulation resources, can improve the experience of end users, who are interested in performing visual analysis on spatiotemporal data.

This paper proposes a new system, called GeoSTAT - GEOgraphic SpatioTemporal Analysis Tool, for visualization and analysis of spatiotemporal data which takes into account, the six essential characteristics discussed by Andrienko et al. [Andrienko et al. 2010b], as mentioned previously. A case study using the GeoSTAT system was proposed to perform a spatiotemporal analysis using data on fire spots and failure events in power transmission lines, aiming at finding evidences that support the hypothesis that fires occurring close to transmission lines could be the cause of failure events in the power system.

The rest of this paper is organized as follows. Section 2 discusses related work. Section 3 focuses on the presentation of the proposed system. Section 4 addresses a case study to validate the proposed ideas. Finally, section 5 concludes the paper and presents further work to be undertaken.

2. Related Work

This section focuses on related works concerning the visualization and analysis of spatiotemporal data.

Ferreira et al. [Ferreira et al. 2011] propose an interactive visualization system that supports the visual analysis of spatiotemporal bird distribution models. It is a spatiotemporal approach towards the specific domain of birds. It is important to highlight that besides being valid for just one specific domain, the solution does not provide mechanisms to connect to external databases, being constrained to the database developed by the authors.

Andrienko et al. [Andrienko et al. 2010a] propose a framework based on the Self Organizing Map technique (SOM) [Kohonen 2001], a combination of clustering and dimensionality reduction. This technique follows the idea that objects are not just

clustered, but also arranged in a space with one or two dimensions, according to their similarity as a function of multidimensional attributes. It is possible to conclude that the use of this technique deals with both spatial and temporal dimensions, allowing coherent analysis of spatiotemporal data. The technique is domain-independent, and seems to be useful in any knowledge field, besides bringing the idea of clustering for aggregating and reducing the database. However, it is important to notice that this work does not provide interoperability between heterogeneous datasets.

Roth et al. [Roth et al. 2010] present a web mapping application that supports spatiotemporal exploration in the criminology domain. The application offers a spatiotemporal browsing resource which animates simultaneously a map and a frequency histogram illustrating the temporal distribution. This application enables the visualization of the variation of data through time, organized into crime categories. Despite this solution supports spatiotemporal data, it is limited to one specific application domain and there is no database interoperability.

Reda et al. [Reda et al. 2009] developed a visual exploration tool to analyze changes in groups in dynamic spatiotemporal social networks. They propose two interesting techniques for spatiotemporal visualization. The *affiliation timeline* displays the structure of the community in the population and its evolution in time, and the *spatiotemporal cube* enables the visualization of the movement of communities in a spatial environment. However, besides being valid only for the domain of social groups, it does not describe how the user should supply the data for visualization and analysis. We conclude this solution has some limitations concerning data heterogeneity.

Andrienko et al. [Andrienko et al. 2007] address a framework for visual analysis of spatiotemporal data representing the trajectory of mobile objects. The framework combines database operations with computational processing, data mining and interactive visual interfaces. This solution highlights the use of the OPTICS clustering algorithm for detection of frequently visited places and database reduction. It is a domain-independent solution, though it is constrained to the trajectory of mobile objects represented by points in space. Besides, the authors do not make clear the acceptable format for the trajectory data.

Among the previously mentioned research works, which focus on the visualization and analysis of spatiotemporal data, some of them address domain-specific solutions, thus being useful for a limited group of users. Furthermore, many of them do not provide flexibility concerning the use of heterogeneous datasets, often requiring a considerable effort from users to adapt their datasets to the chosen application in order to perform the analysis.

There are also problems concerning usability, as the user interfaces do not provide to end users enough freedom to include or remove feature types that they might find relevant to their tasks.

3. The Geographic Spatiotemporal Analysis Tool

This section introduces GeoSTAT (Geographic Spatiotemporal Analysis Tool), a new web-based system for spatiotemporal visualization and analysis.

Through the GeoSTAT system, the user interested in viewing and analyzing a spatiotemporal dataset will be able to use several visualization resources that deal with

both spatial and temporal dimensions. Besides, clustering-based data mining algorithms, adapted for the spatiotemporal domain, were integrated into the system. Besides the advantages of being a web application, GeoSTAT was conceived under the generality point of view. For this reason, it is a domain-independent system, which can be connected to any spatiotemporal data source available over the Web by implementing the spatial data sharing services specified and standardized by the OGC (Open Geospatial Consortium) [OCG 2011].

3.1.Components

The interactive user interface of GeoSTAT system is comprised of ten components responsible for the functionalities offered by the system. Figure 1 presents this interface and enumerates these components: 1) map; 2) spatiotemporal layers (overlap); 3) temporal controller; 4) temporal filter; 5) spatial filter; 6) temporal distribution graphic; 7) data mining results; 8) actions menu; 9) data mining; 10) information about the connected data servers.

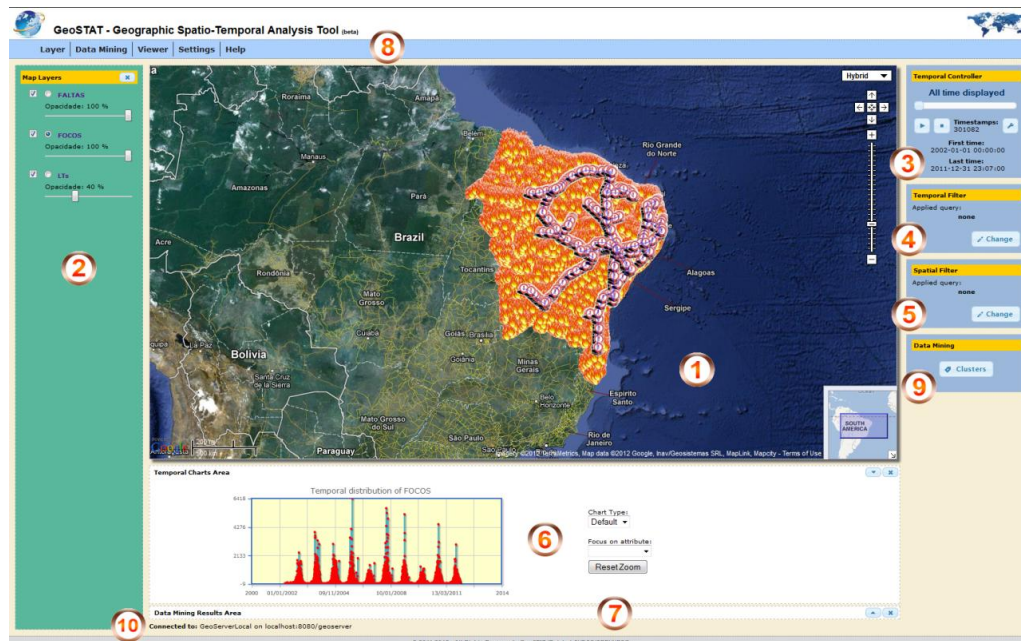


Figure 1. The main interface and components of GeoSTAT system displaying data layers used in case study presented in section 4.

The map component uses the Google Maps API to offer a dynamic map. The spatiotemporal layers component allows users to add layers and spatiotemporal (or just spatial) data published in servers that implement the OGC WMS (Web Map Service) and WFS (Web Feature Service) services. These data are plotted on the map, and made available through the components that deal with the temporal dimension, such as the temporal controller, the temporal filter and the temporal distribution graphic. They are also made available for clustering-based data mining through the system.

Through the use of the temporal controller, it is possible to change the map visualization using a temporal filter. This filter can be defined as either a given instant (timestamp), or a more abstract level of temporal resolution, such as months, for example. The temporal controller also allows the production of a temporal animation,

which lets the user to visualize on the map the eventual changes in the spatial distribution of the data as a function of the temporal variation. It also displays a specific timestamp and enables the observation on the map of a spatial distribution of data on this timestamp. Still, it may terminate the animation and view the spatial distribution of the whole dataset on the map again, regardless of the temporal dimension.

Besides the temporal controller, another available temporal visualization resource is the temporal distribution graphic. It is responsible for helping the user to visualize changes in the spatiotemporal data as a function of time, adding to the map resource, which helps the visualization of the distribution as a function of space.

The spatial and temporal filter components are responsible for the spatial and temporal query and selection, respectively, of the data visualized through the spatiotemporal layers. Through the temporal filter, the user may, by means of four filter options and observing the temporal resolution used, reduce the spatiotemporal dataset for visualization and analysis. The four options available for the temporal filter are: *from*, *until*, *in* and *between*. On the other hand, through the spatial filter, it is possible to visualize a topological relationship between two spatial or spatiotemporal layers previously added to the system, regardless of the source data source. It is possible to perform the following topological relations between two layers: *intersects*, *contains*, *crosses*, *touches*, *covers* and *overlaps*. It is also possible to apply negation (*not*) to each one of these relations, in cases where this is relevant for the analysis performed by the user.

In the component of data mining, it is possible to perform the clustering-based data mining in the previously added layers, view the result of a previous data mining process and the detailed status of data mining processes under execution. The data mining processes run in background, so users do not need to wait for the end of this processing, as they may perform other tasks.

The component of data mining results is responsible for offering the statements necessary for the spatiotemporal visualization and for browsing a layer containing data mining results. The user may browse through the timestamps that have the occurrence of clusters and view each cluster separately on the map. If the data mining is made with two layers, the user will have the option of viewing just the relevant clusters, that is, those which have at least one point of each layer, as well as options to view just the clusters that group only points of one layer. It is also possible to see all clusters of a given timestamp, or even all clusters.

Finally, the actions menu component offers shortcuts for the rest of the components of the interactive graphic interface of the GeoSTAT system, and the connected source data server component is responsible for displaying information about the data servers that are connected to a user session of the system.

3.2. Architecture

The GeoSTAT system architecture is defined using three layers: visualization, control and persistence.

The visualization layer is responsible for the user interface, offering components for loading, handling and visualizing the data through the temporal and spatial dimensions, presented in section 3.1.

The control layer is responsible for the processing of all requests generated and sent from the visualization layer, besides being responsible for the communication with the persistence layer, therefore being the kernel of the GeoSTAT system. Figure 2 presents the five existing modules in the control layer. These modules are activated according to the nature of the request to be processed by this layer.

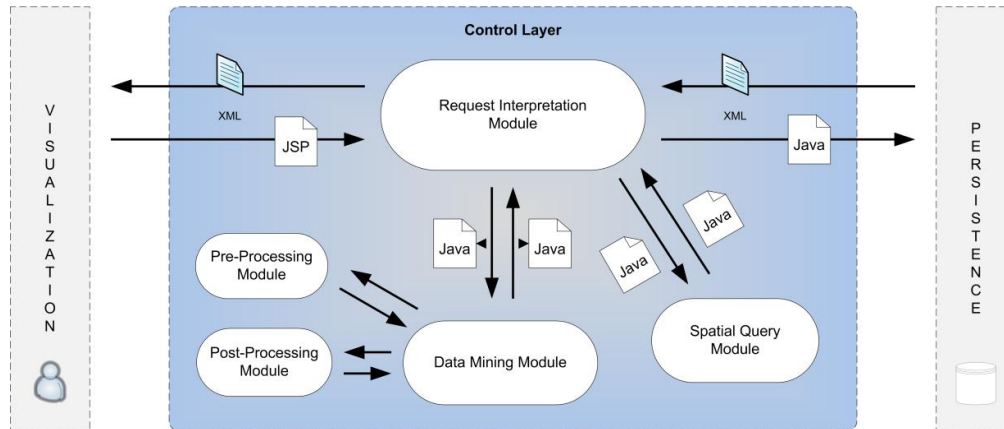


Figure 2. Control modules of the GeoSTAT system architecture.

The request interpretation module (see Figure 2) is the main module of the control layer. It is responsible for receiving and treating every request coming from the visualization layer and for establishing contact with the other modules, besides making contact with the persistence layer. There are two types of treatment to the requests that arrive at the request interpretation module: query or data delivery requests and data processing requests, that is, data mining or spatial query requests. The data requests are sent directly to the persistence layer, which is responsible for interpreting and processing this kind of request. On the other hand, the data processing requests may be forwarded to the data mining module or to the spatial query module.

The spatial query module (see Figure 2) is responsible for the processing of spatial queries between two different layers. The result of the query processing (spatial filter) is sent to the visualization layer, for exhibition to the end user.

The data mining module integrates several known clustering algorithms. These algorithms were obtained from the Weka toolkit [Hall et al. 2009]. Seven algorithms were adapted and are available on the GeoSTAT system: COBWEB, DBScan, K-Means, X-Means, Expectation-Maximization, Farthest-First and OPTICS. Hence, GeoSTAT system is capable of performing clustering-based spatiotemporal data mining on any spatial or spatiotemporal database. The output returned by the data mining module is stored in a spatiotemporal database and made available for query from the system, as soon as the processing is complete. The data mining module uses threads for concurrent processing.

In order to make possible the spatiotemporal integration and adaption of the several data mining algorithms used, we developed the data pre-processing and post-processing modules. These auxiliary modules are responsible for preparing data to be used by the algorithm selected by the user and preparing the results obtained through the execution of this algorithm for treatment by the visualization layer, respectively.

The persistence layer is responsible for connecting the GeoSTAT system to the databases requested by the users through the components of the visualization layer. When a data request is received from the control layer, the persistence layer first identifies the type of connection that will be established. It can connect either to the OGC WMS and WFS services, or to a spatiotemporal database developed to operate exclusively with the system. The OGC services are accessed from their web servers.

The spatiotemporal database stores information used by the GeoSTAT system to connect to the OGC services, as well as the complete results of the data mining processes performed by the system and available for visualization.

4. Case study: Analysis of spatiotemporal correlation between failures in power transmission lines and fire spots

This study consists in the analysis of two sets of spatiotemporal data. Each set is comprised of records of an spatiotemporal event.

4.1. Data

To carry out this study, we used georeferenced spatiotemporal data about fire spots detected in the Northeastern region of Brazil, supplied by the National Institute for Space Research¹ (INPE), through the Weather Forecast and Climatic Studies Center (CPTEC), which publishes this kind of information daily, through their Fires Monitoring Portal².

We obtained a total of 2,361,040 records of fire spots detected in the region, in the period between 01-01-2002 and 12-31-2012, that is, in the last ten years. The spatiotemporal data were obtained in the ESRI™ Shapefile format, using the WGS84 geographic reference system, and temporal data according to the GMT. According to INPE, their system detects the presence of fire in the vegetation and the mean error in the spatial location of the spots is of approximately 400 meters, with standard deviation of about 3 kilometers, and with about 80% of the spots detected in a distance of one kilometer from the coordinates indicated by the system. In the temporal validity, the satellites offer a mean temporal resolution of 3 hours. This is the mean time between the pass of two satellites capturing information about the same region.

Another spatiotemporal database was used in this study. It is about failure events in power transmission lines, recorded by the San Francisco Hydroelectric Company (Eletrobrás/Chesf), which operates throughout the Northeastern region of Brazil. Since we could not get official data from Eletrobrás/Chesf, due to technical and confidentiality matters, we developed an algorithm to generate spatiotemporal failure events randomly, obeying the spatial constraint imposed by Eletrobrás/Chesf's transmission line network, and the temporal constraint imposed by the other database used in this study.

We generated a total of 131,834 failure records in Eletrobrás/Chesf's transmission lines, in the period between 01-01-2002 and 12-31-2012, that is, also in the last 10 years. These records were stored in a spatiotemporal database, also in the WGS84 geographic reference system, and with temporal information according to the

¹INPE –Brazilian National Institute for Space Research. More information at <http://www.inpe.br/>

²INPE/CPTEC –Fires Monitoring Portal. Available at <http://www.inpe.br/queimadas/>

GMT. Aiming at helping in the visual analysis of the transmission line failure events, we also used a set of spatial data containing Eletrobrás/Chesf's transmission line network.

Both datasets used in this study share the same spatial geometry (POINT) and also the same temporal resolution (timestamp). In order to use the data in GeoSTAT system, we needed to install Geoserver web map server and create layers for each dataset.

To conduct this study, the GeoSTAT system user will be called analyst, a specialist user in the approached domain, looking for relevant information implicit in a large volume of spatiotemporal data.

4.3. Experiment

Figure 1 shows the GeoSTAT system interface with the three spatiotemporal layers loaded into the system from the data connection with Geoserver. What is seen is the result of about two million and a half points plotted in the map, enough to fill the whole Northeastern region.

The temporal distribution graphics, generated and shown automatically when a spatiotemporal layer is loaded and selected in the GeoSTAT system, allows the analyst to verify the behavior of the whole volume of data. By observing the graphic corresponding to the fire spots layer (showed in Figure 1), we notice that there is an annual repetition of the distribution of the number of spots detected, where the maximums concentrate in the first and in the last months of each year. This is the period when the Northeastern region registers the highest temperatures, which contributes to the occurrence of new fire spots. Through this graphic, we can also observe that the maximum number of spots detected in one day, in the 10-year period, was of 6,418 spots. This number was reached in 11-07-2005.

By observing the graphic corresponding to the transmission line failures layer, we notice a temporal behavior that is practically continuous. Once the data was randomly generated through an algorithm, the temporal distribution of the occurrences was uniform, registering the maximum of three occurrences in one single day.

For a better visualization of the power line failures and of the detected fire spots, it might use a more generic temporal resolution than timestamp, such as "Date and Time", for example, joining all the records occurring between "10-15-2011 15:00:00" and "10-15-2011 15:59:59" in one single view, for example. This strategy allows several simultaneous visualizations, time-time, of failures and fire spots within 10 years of data. However, the cost would be too high for the analyst to view image by image, time by time, manually, to find interesting behaviors. The use of the clustering technique emerges as a good option to reduce the cost to the analyst, by making the spatiotemporal clustering of the events.

With the layers "FAILURES" and "SPOTS" added to the GeoSTAT system, we activate the spatiotemporal clustering option offered by the system to perform the data mining with both layers. This option enables the analyst to view the spatiotemporal clusters of each separate event and the relevant clusters, that is, the spatiotemporal clusters containing records of both events.

In order to execute the data mining, besides the three input layers, the user had to inform the following required parameters: “Date + 3-3 hours” for temporal resolution, and DBScan as the data mining algorithm, with MinPoints = 2 and Epsilon = 0.013472.

The choice of the value 0.013472 for the Epsilon parameter of DBScan is due fact that one second (angular measurement unit) is approximately equal to 30.9 meters. Since about 80% of the fire spots detected by INPE occur within one kilometer from the indicated coordinates, and the mean error in the spatial location of the records is of 400 meters, we thought reasonable that the radius of a generated cluster ranged from 1 to 1.5 kilometers. Since 48.5 seconds is approximately equal to 1,498.65 meters (1.5 kilometers) and one decimal degree has 60 minutes and 60 seconds, then we conclude that 1,498.65 meters is approximately equal to 0.013472 meters.

4.4.Results and Conclusions

The data mining process of this case study lasted 7 hours, 37 minutes and 5 seconds. It was executed in a web application server, running Microsoft™ Windows 7 Professional (64-bit) operating system, with Intel™ Core i7 processor and 16 GB of RAM.

The statistical results for the classification of the records after the execution of the algorithm showed that only 32,275 records, 1.29 of the whole dataset, were considered relevant by the GeoSTAT system. This means that only these records are contained in relevant spatiotemporal clusters, those which contain records of both studied events. Approximately 86.03% of the records were associated to a spatiotemporal cluster. The rest of the records, 13.97% of the total, were considered outliers because they do not belong to any spatiotemporal cluster, representing only isolated occurrences in space-time.

From the 318,901 spatiotemporal clusters generated, just 1,376 (0.43%) were considered relevant under the viewpoint of the measurement parameters used in the execution of the data mining algorithm. Each irrelevant cluster grouped, on average, 6,623 records, while each relevant cluster grouped, on average, just 23 records.

Figure 3 presents a screenshot captured from the GeoSTAT system showing in the map all the relevant spatiotemporal clusters generated for the 10-year period of the dataset. The first information that may be noticed by the analyst in this visualization is that the region which concentrated more clusters was the region located in the Southeast of the state of Ceará, more precisely in the border with the states of Paraíba and Rio Grande do Norte, highlighted in the picture. The metropolitan regions of Maceió-AL and of Recife-PE, as well as the region of the city of Sobral-CE, are also regions with many clusters.

The generated spatiotemporal clusters can be browsed with the components for temporal selection and, from this definition, with the individual selection of each cluster corresponding to the previously selected timestamp. The analyst may choose the visualization of relevant clusters only, or the visualization of all clusters. The analyst may also visualize each individual cluster, or visualize all the clusters, regardless of the temporal dimension.

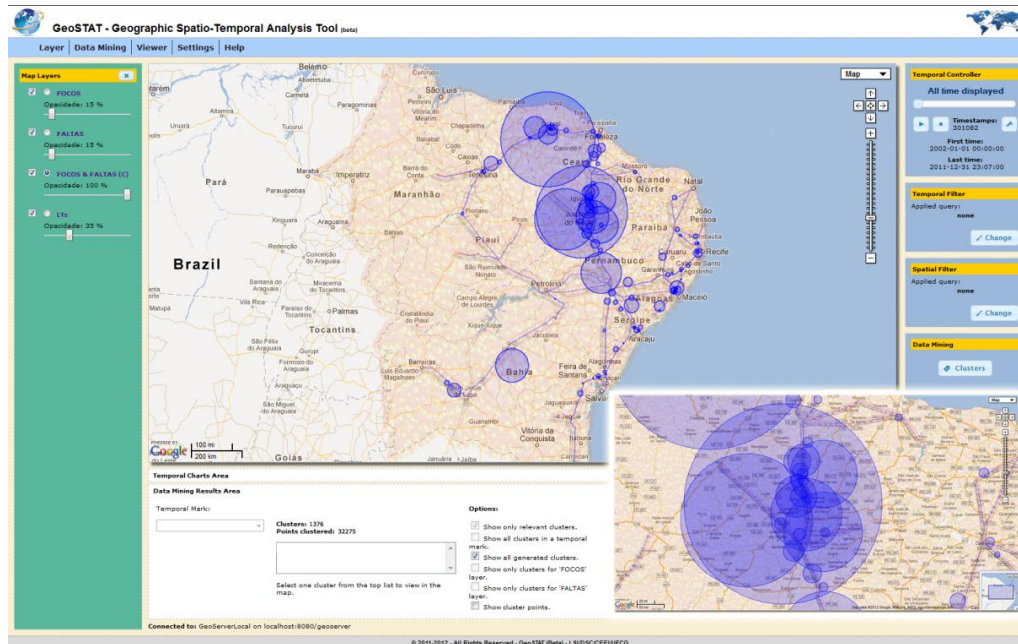


Figure 3. GeoSTAT system showing all the relevant clusters.

For the analyst, interested in confirming the hypothesis that some fire spots are the cause of failures in power transmission lines, Figure 4 exemplifies a case where the hypothesis is confirmed. A failure occurring in the line “FORTALEZA II - CAUIPE” at 03:14 p.m. in 11-03-2004 had its cause pointed as “FIRE” and, besides, due to the data mining performed together with data from records of fire spots detected in that region at the same period as the failure, pointed out a spatiotemporal clustering between this failure and two fire spots: one detected at 04:08 p.m., with approximate distance of 1 kilometer from the failure, in East direction; and another one, detected at 04:01 p.m., with approximate distance of 1.5 kilometers from the failure, in the North direction. If we consider the spatial precision errors and the temporal resolution of these data, the analyst could point these two fire spots as the actual causes of the failure.

The results achieved with the use of the GeoSTAT system were satisfactory for the application domain explored in this study. The visualization resources explored allowed the discovery of interesting implicit information, from two large volumes of data.

It is important to observe that the statistical data mining results pointed to an index of relevant clusters under what most specialists in this kind of event would expect. This is due, mainly, to the use of simulated records of power transmission line failures. The use of real data, captured and structured by Eletrobrás/Chesf will certainly produce better results, as the presence of more relevant clusters.

Besides using real data, the specialists have, through the GeoSTAT system, several spatiotemporal clustering algorithms available. Their results may be compared and analyzed to find new relevant information.

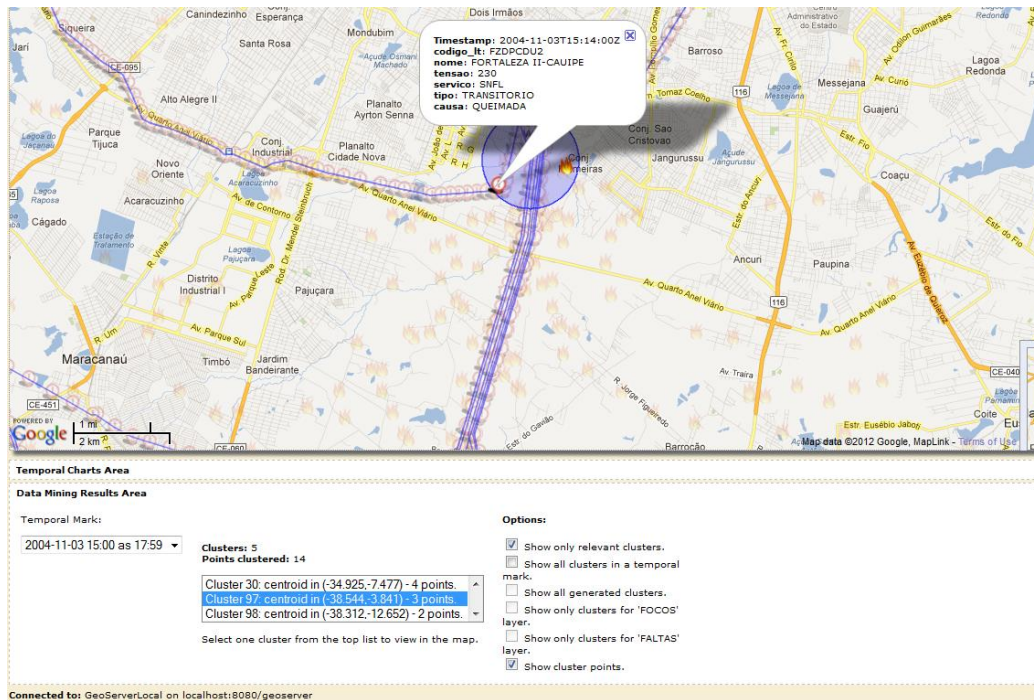


Figure 4. GeoSTAT system displaying, in detail, the spatiotemporal cluster no. 97, with temporal mark “11-03-2004 03:00 p.m. to 05:59 p.m.”.

5. Conclusion and Future Work

In this paper, we proposed a system for visualization and analysis of spatiotemporal data. This system managed to address the six features needed by a solution for spatiotemporal visualization and analysis: resources for the spatial dimension, resources for the temporal dimension, domain independence, flexibility, interoperability and data mining based on spatiotemporal clustering. It is a solution that prioritizes the end user, offering a set of functionalities that allow the execution of a job, in a practical and efficient manner.

Finally, we conclude that the proposed system met its objectives, proving to be satisfactory and efficient. We also conclude that many improvement issues can be addressed in future studies, which certainly will contribute to a more robust system. One point is the inclusion of another data mining technique such as spatiotemporal association rules.

References

- Andrienko, G., Andrienko, N. & Wrobel, S. (2007), “Visual Analytics Tools for Analysis of Movement Data”, SIGKDD Explorations 9(2), 38–46.
- Andrienko, G., Andrienko, N., Bremm, S., Schreck, T., Landesberger, T. V., Bak, P. & Keim, D. (2010a), “Space-in-Time and Time-in-Space Self-Organizing Maps for Exploring Spatiotemporal Patterns”, Computer Graphics Forum 29, 913–922.
- Andrienko, G., Andrienko, N., Demsarb, U., Dranschc, D., Dykes, J., Fabrikant, S. I., Jernf, M., Kraakg, M.-H., Schumannh, H. & Tominskih, C. (2010b), “Space, time

- and visual analytics”, *International Journal of Geographical Information Science* 24(10), 1577–1600.
- Andrienko, N., Andrienko, G. & Gatalsky, P. (2003), “Exploratory spatio-temporal visualization: an analytical review”, *Journal of Visual Languages & Computing* 14(6), 503–541.
- Bédard, Y., Merrett, T. & Han, J. (2001), “Fundamentals of Spatial Data Warehousing for Geographic Knowledge Discovery”, Vol. *Research Monographs in GIS*, Taylor & Francis, chapter 3, pp. 53–73.
- Ferreira, N., Lins, L., Fink, D., Kelling, S., Wood, C., Freire, J. & Silva, C. (2011), “BirdVis: Visualizing and Understanding Bird Populations”, *IEEE Transactions on Visualization and Computer Graphics* 17(12), 2374–2383.
- Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P. & Witten, I. H. (2009), “The WEKA data mining software: an update”, *SIGKDD Explorations* 11(1), 10–18.
- Johnston, W. L. (2001), “Information visualization in data mining and knowledge discovery”, Morgan Kaufmann Publishers, San Francisco, CA, USA, chapter 16. *Model Visualization*, pp. 223–227.
- Kohonen, T. (2001), “Self-Organizing Maps”, Vol. 30 of *Information Sciences*, 3rd ed., Springer-Verlag.
- Kopanakis, I. and Theodoulidis, B. (2003), “Visual data mining modeling techniques for the visualization of mining outcomes”, *Journal of Visual Languages and Computing* 14(6), 543–589.
- OGC (2011), “OGC - Making Location count”, Open Geospatial Consortium. Disponível em: <http://www.opengeospatial.org/>.
- Reda, K., Tantipathananandh, C., Berger-Wolf, T., Leigh, J. & Johnson, A. E. (2009), “SocioScape - a Tool for Interactive Exploration of Spatiotemporal Group Dynamics in Social Networks”, In: *Proceedings of the IEEE Information Visualization Conference (INFOVIS'09)*, Atlantic City, NJ, USA, pp. 1–2.
- Roth, R. E., Ross, K. S., Finch, B. G., Luo, W. & MacEachren, A. M. (2010), “A user centered approach for designing and developing spatiotemporal crime analysis tools”, In: *Proceedings of GIScience 2010*, Zurich, Suíça, pp. 66–71.