CO-422

MULTIPLE-VIEW SYSTEM FOR SOCIAL NETWORK ANALYSIS AND EXPLORATION

DELAZARI L., BRANDALIZE M.C.B., SLUTER C.R.
Parana Federal University, CURITIBA, BRAZIL

ABSTRACT

A proposal for a multiple views system for cartographic representation of a social network is presented in this paper. Considering that social network analysis need to be based on graphs, the main issue is how to establish a digital data structure and thematic representation for these networks in order to preserve the spatial positions and relations of the network actors. The proposed solution for this research problem is to combine different representations of the phenomena in a computational system with multiple views: a traditional thematic map at a suitable scale depicting a whole network whose actors are located in a Brazilian big city; the correspondent graph, tables and charts. The system allows users to explore data and representations, to interact with them and provides interactively linked views.

1. BACKGROUND AND OBJECTIVES

This paper describes the design and implementation of a multiple views system for social network analysis. Social networks consist of individuals, groups of people or institutions, called actors that are connected to each other. Commonly, social scientists use graphs and matrices in order to analyze these networks. Graphs are formed by nodes and arcs, which represent, respectively, network actors and their relationships. Graphs representations or matrices structures, however, do not take into account the spatial location of actors and their relationships. Consequently, the network attributes are not spatially represented and so it is not possible to analyze and understand their spatial structure.

The use of graphs as social network representation has two known problems: (1) nodes are not spatially located and (2) symbols for nodes and arcs are not suitable to represent attributes of the graphs elements. The representation of spatial location is needed to analyze actors' proximity and neighborhood relationships. Moreover, the attributes (characteristics) of the graphs elements must be depicted in order to make possible to analyze clusters, dispersions, tendencies and influence areas.

These problems have been solved by the map design for social networks presented for papers and for computer screens. The use of thematic maps allowed the social scientists to understand the dynamics of the network relations and consequently the existence of regions characterized by cultural, political and economical influence. However, after some experiences with these thematic maps, the social scientists noticed the need for tools which would provide geovisualization capabilities for analyzing graphs, tables and maps by a set of linked views. Some of those needs are:

a) To identify actors' proximity and neighborhood relationships which are possible through cartographic visualization;

b) To identify actors' concentration in specific regions, what can lead to important conclusions about why some geographic regions are better assisted than others, and consequently, to plan the location of some new actors in regions without social assistance;

c) To identify the actors' characteristics related to different geographic locations and which regions are supported by governmental or non-governmental organizations, or both;

d) To visualize the distances and orientations between actors based on the geographic representation of their links;

e) To identify the geographic concentration of the network actors related to different geographic levels: municipal, state, national or even worldwide;

f) To compare the distribution of actors on graphs and thematic maps in order to identify the concentration in both representations;

g) To have access to actors and links attributes.

When the knowledge on thematic mapping and geovisualization is applied to cartographic representation of social networks, it is possible to develop data exploration in order to know about the proximities of actors and their neighboring relations. Moreover, the thematic mapping of actors’ attributes allows the analysis of clustering and diffusion, tendencies, regions of influence, and so on. The importance of spatial analysis for social network studies leads to our main research objective which is to propose a system with multiple linked views. This system provides ways to dynamically manipulate and explore different graphic representations.

2. APPROACH AND METHODS

2.1 Social Network Analysis

Social networks are difficult to visualize, navigate, and analyze and it is considerably difficult to find relevant patterns on networks. Network analysts focus on relationships instead of the individual elements which can explain social, cultural, or economic phenomena, but how the elements are connected is just as important as the elements themselves. Using newer techniques analysts can find patterns in the structure and learn how individuals are influenced by their surroundings (PERER, 2008).

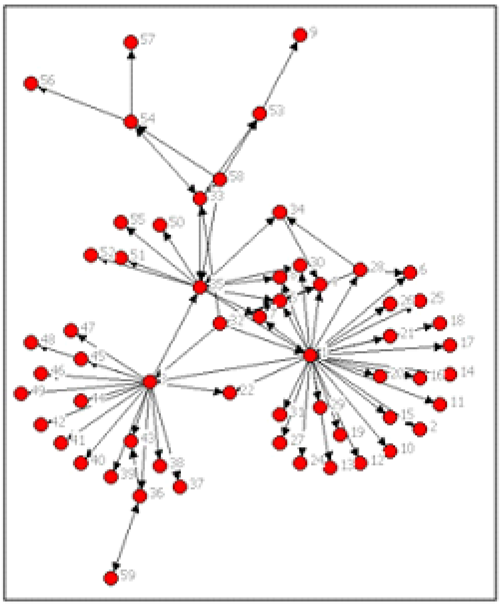
Using visualizations to assist in Social Network Analysis (SNA) is not a new concept for sociologists. Visual images can be used to examine the patterns of network data, as described in Freeman (2004). A history of the use of visual images in social networks is described in Freeman (2000) and includes one of the earliest known examples of social network visualization from Jacob Moreno in1934.

Usually social scientists use two different ways to represent social networks: graphs and matrices (Figure 1). These two forms of representation are generated by using statistical and visualization techniques software for network analysis. Two of the most popular tools used by social scientists are UCINET and Pajek. Each of these tools has a set of feature to measure social networks, grounded in the theory and techniques of sociologists.

UCINET probably is the most known and frequently used SNA software. It contains a large set of features applied to networks analysis as centrality, proximity and so on. The software also allows graphic functions to generate scatterplots, dendograms and trees (http://www.analytictech.com/ucinet/).

Pajek, accordingly to its authors (http://vlado.fmf.uni-lj.si/pub/networks/doc/pajek.pdf) has as main goals: to support abstraction by factorization of a large network into several smaller networks that can be treated further using more sophisticated methods; to provide the user with some powerful visualization tools; and to implement a selection of efficient algorithms for analysis of large networks.

Although both software have graphical capabilities that allow change in shape, color and size of networks, there's no spatial reference associated to any kind of process.



	Actors	1	2	3	4	5	6	7
1	FAS	0	1	1	1	1	1	0
2	COMTIBA	0	0	0	0	0	0	1
3	CMFAD	0	0	0	0	0	0	0
4	CMDPPD	0	0	0	0	0	0	0
5	COGEMAS	0	0	0	0	0	0	0
6	CMAS	0	0	0	0	0	0	0
7	CASDM	0	0	0	0	0	0	0
…								

Figure 1 – Graph and table related to a network

Another software developed to help social scientists is called SocialAction, which integrates statistical and visualization techniques (Figure 2) (PERER, 2008) but also does not employ any spatial reference of actors. SocialAction embeds statistical algorithms to detect important individuals, relationships, and clusters. Instead of presenting statistical results in a typical tabular form, results are integrated as a network's visualization while providing meaningful computed attributes of the nodes and edges. Along with computed attributes, users can easily and dynamically filter nodes and edges to discover interesting data characteristics (PERER, 2008).

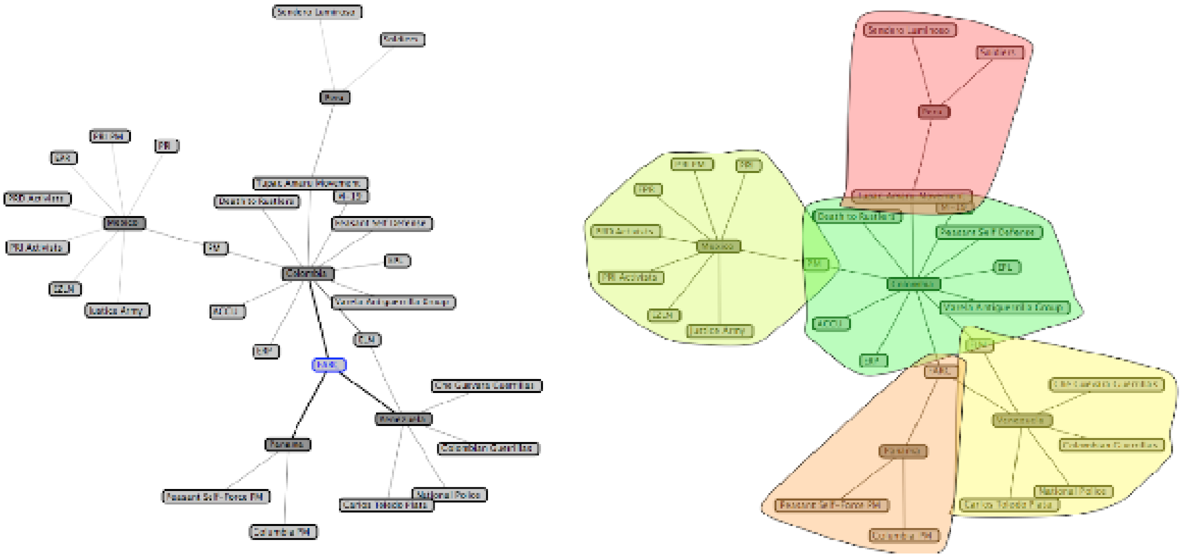


Figure 2 – (a) SocialAction network representation; (b) Colored polygons surround the subgroups of the network.

Source: Perer and Shneiderman (2006)

Social Network Analysis is similar to geospatial data analysis since both are made through visual map analysis and queries in a Geographic Information System (GIS). Therefore graphic solutions used in social networks representations must be prepared in accordance to technical-scientific knowledge of Cartography science.

3. MULTIPLE VIEW SYSTEM - THEORETICAL CONSIDERATIONS

In 1990 Dibiase proposed a visualization model of the role of maps in scientific visualization. In this model, map functions are related to stages of analysis and planning processes, which are exploration, confirmation, synthesis and presentation. An exploratory geovisualization system allows users to investigate possibilities and test scenarios. In the exploration stage a user can change visualization parameters, use different views in order to make comparisons with different parameters or compare different representations of the same data, as maps, graphics and tables. In this case, those views must be linked for allowing the user to analyze the data by using combined navigation or brushing, as presented in Figure 3.

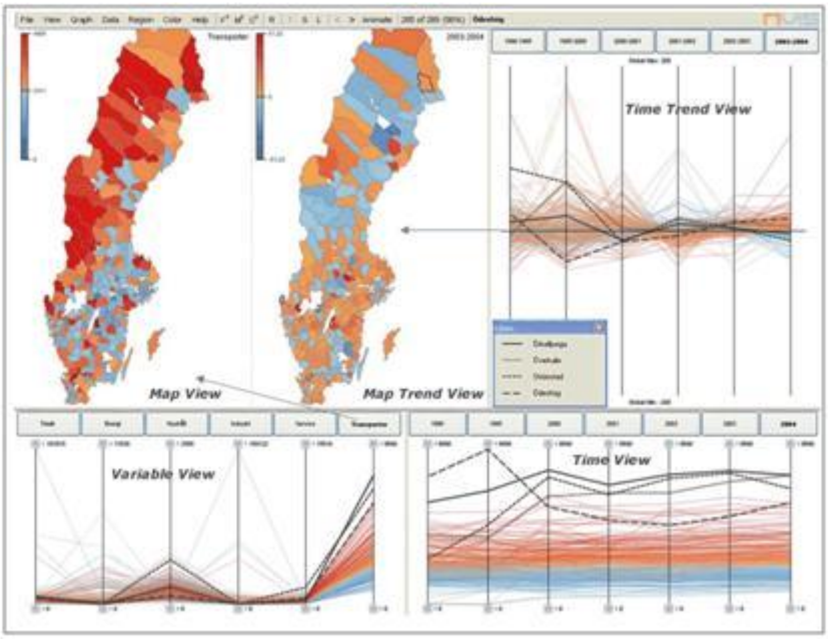


Figure 3 – GAV Multiple View System (GeoAnalytics Visualization)

Source: Jern et al (2007)

In Figure 3 one of the views represents spatial data and the others are parallel coordinate graphics that present several variables simultaneously, allowing comparisons. When users point out a polygon in the map, the line correspondent to each one of the variables in those parallel coordinate is highlighted.

There are others systems with the same functionalities. For example, GEOVIZ tool, developed in “GeoVista Center”, at Pennsylvania University, enables multidimensional geographic data visualization. Its components are related meaning selecting a set of variables will reflect in further components in use. GEOVIZ is used for spatial, temporal and attributes analysis. The application focus is potentially health, socioeconomic and demographic data (ROBERTS, 2007).

The CommonGIS project was originated from IRIS (Information Retrieval Intelligent System), developed in C++ in the 1990s. Later it was renamed Descartes, an internet system developed in Java, and finally, in 1999 became CommonGIS. Its functions include remote access through internet, manipulation, exploration and data analysis tools, automatic map generation, and maps modification using direct data manipulation (ANDRIENKO e ANDRIENKO, 2003).

It can be observed that systems found in the literature are developed mostly focusing on health and demographic applications. Almost all of them use statistical data aggregated over previously defined regions (states, councils, census sectors), and remaining data are scatterplots, parallel coordinates and tables.

The peculiarity of social network data constitutes a new proposition in the context of multiple view systems because it has been verified the need of an analysis through graphs and maps. On the other hand, software developed for network analysis does not work with spatial data reference.

4. MULTIPLE VIEW SYSTEM PROPOSITION

4.1 Social Network Spatial Representation

The social networks that have been studied so far take place in Curitiba (KAUCHAKJE, 2007; DELAZARI, 2008, BRANDALIZE, 2009), and these studies include sociology, map design, cyberspace and cartography. Such researches were developed employing empirical techniques comprising the following steps:

- Data collection: user interviews, document analysis, and internet content exploration;
- Design and presentation of the social networks representation (spatial and non spatial) using UCINET, TOUCHGRAPH, ARCGIS and GOOGLE EARTH;
- Analysis of the results and evaluation of the cartographic process efficiency for the proposed representation through users’ interviews.

The referred methodology was applied to represent and analyze a series of social networks regarding individual’s social rights in Brazil. According the Brazilian Constitution of 1988, the individual’s social rights comprise Housing, Health, Education, Labor, Food Security and Social Assistance. The data necessary to establish these networks were collected through interviews, questionnaires and the Internet. Curitiba, capital of Parana State, Brazil, was used as the geographic main location for the establishment of the networks analyzed in this study. Concerning the referred networks, the main social agents of each network, their partnerships and links were identified, characterized and mapped, first using conventional social network software, like UCINET. Afterwards, the results from these applications were exported to ARCGIS software to proceed with the geographic analyses. The cartographic language was applied to the representation of networks (social agents and their connections) based on the organizational characteristics of the social agents (governmental, non-governmental, third sector) and their connections (if ideological, by project, and so on) (DELAZARI; KAUCHAKJE; PENNA, 2005) (BRANDALIZE, 2009).

The resulted map representations were presented to users in order to verify its efficiency. Users were asked to analyze representations taking into consideration the following points:

a) Is it possible to identify actor’s neighborhood or proximity from network spatial location?

b) Is it possible to identify actors’ classification (governmental, non-governmental, companies)?

c) Is it possible to identify clusters of actors belonging to different levels of government (city, state, federal)?

d) Is it possible to identify links direction?

4.2 Multiple View System Design

Using networks representations, the users were capable of understanding some geographic aspects of them. However, after some experiences with the maps, it was noticed that the users need extra functionalities that allow more data exploration, like maps, graphs and tables in connected views.

The proposed system is being developed in Java, using NetBeans 6.9 and GeoTools. The NetBeans project consists of an open-source Integrated Development Environment (IDE) and an application platform that enable developers to create web, enterprise, desktop, and mobile applications using the Java platform. GeoTools is an open-source Java library that provides tools for geospatial data.

Considering users inexperience with software usage, it was decided to design a clear interface in order to avoid misunderstandings and incorrect choices. Thus, the first research task comprised the system conceptual model design which includes the interface design and all the necessary tools that permit to the user interactively handle maps, graphs and tables as well as visually analyze their relationships.

The result of the interface design comprises two display panels: an interactive map and the correspondent graph, both linked to each other, as presented in Figure 4. There are two menus: File and Network. The menu File enables to open the spatial database that are Curitiba neighborhood limits, a layer of points that represent actors and a layer of lines representing connections. As mentioned previously, the data was collected regarding six different social rights: Housing, Health, Education, Labor, Food Security and Social Assistance. Therefore, there are six different layers of points and lines, one for each theme mentioned.

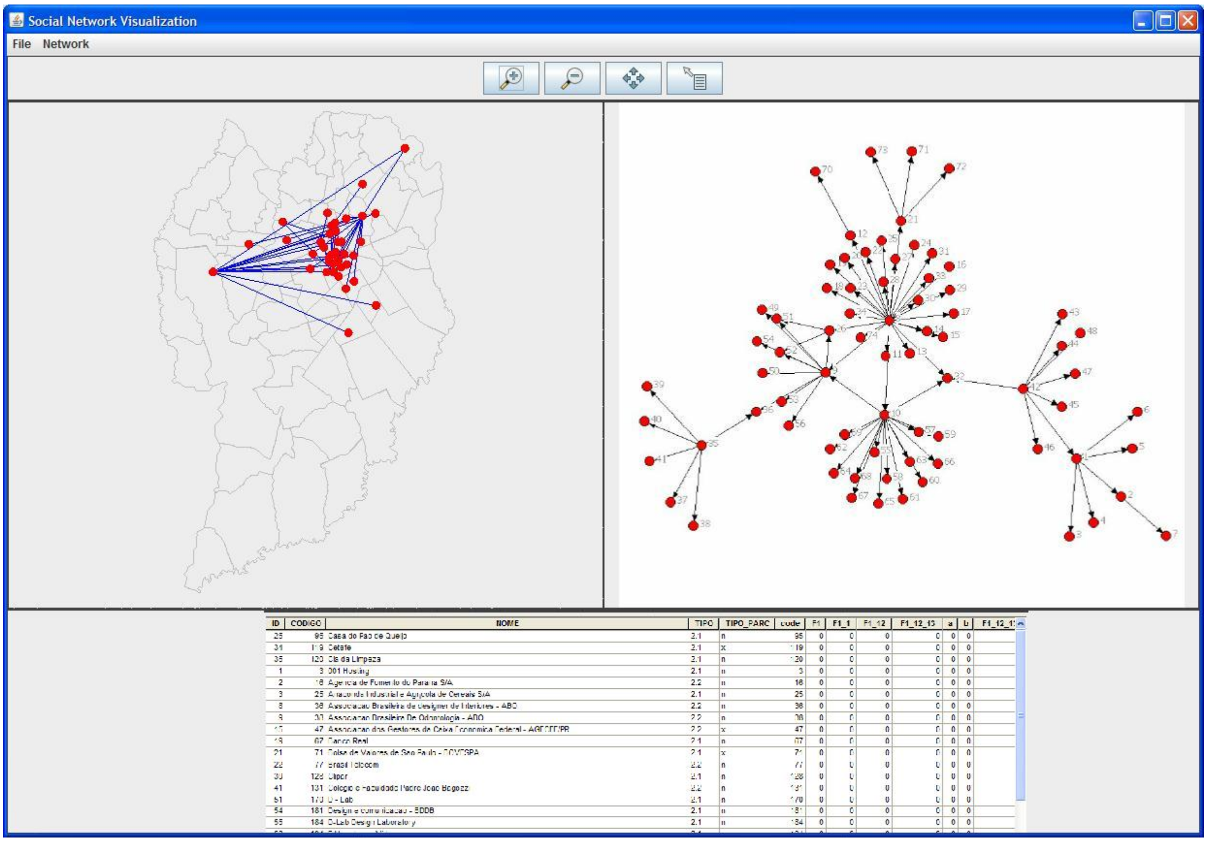


Figure 4 – Initial Interface for Multiple View System

The Menu Network (Figure 5) allows opening the corresponding social right graph. Each actor in the graph has a code that is the same code used in the spatial representation. When an actor is selected, in the map for instance, the correspondent actor in the graph is highlighted, and vice-versa, when a graph's node is selected the matched element in the map is also highlighted.

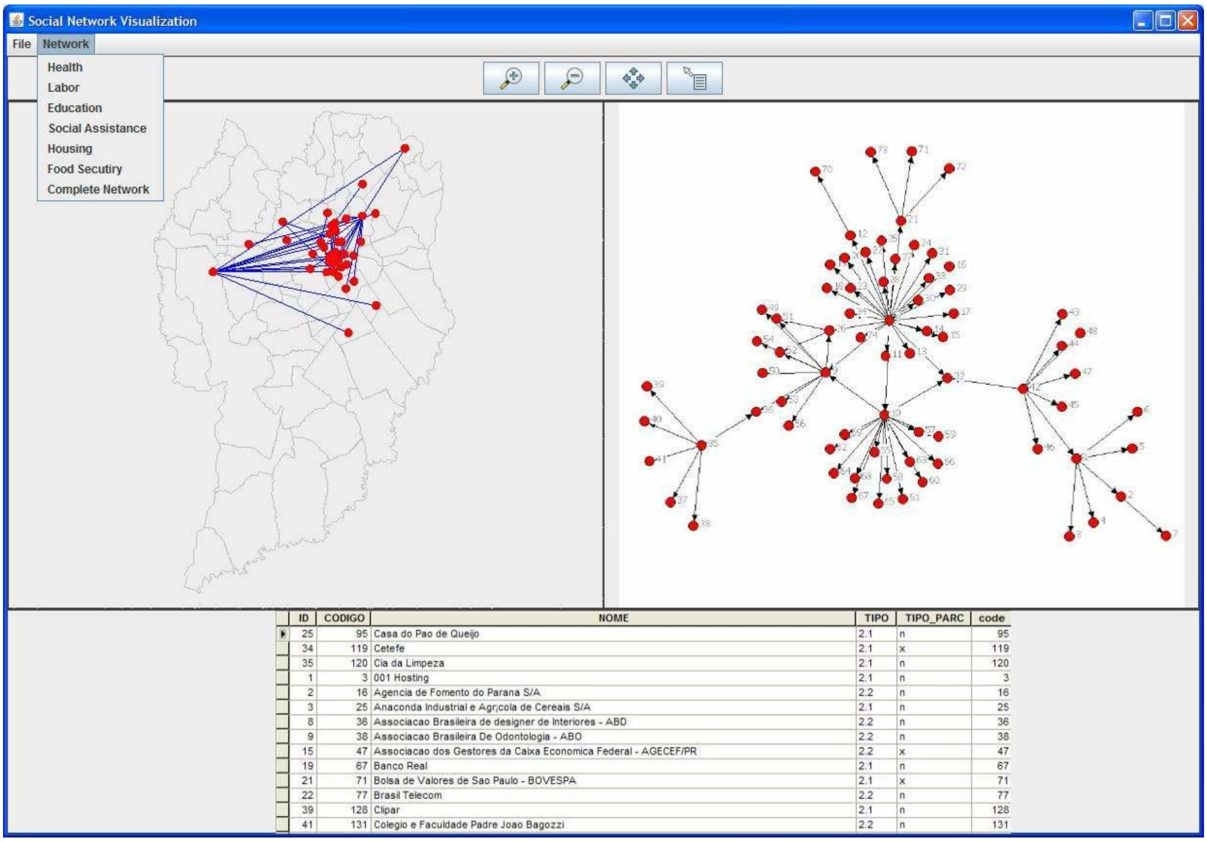


Figure 5 – Detail of Network Menu

Navigation and information buttons allows necessary tools for accessing the details of actors and their connections. Navigation buttons are used to navigate both in the map as in the graph. Information button is used to point out an element in the map or graph and its correspondent elements in the graph and table (at the bottom in the interface) are highlighted (Figure 6).

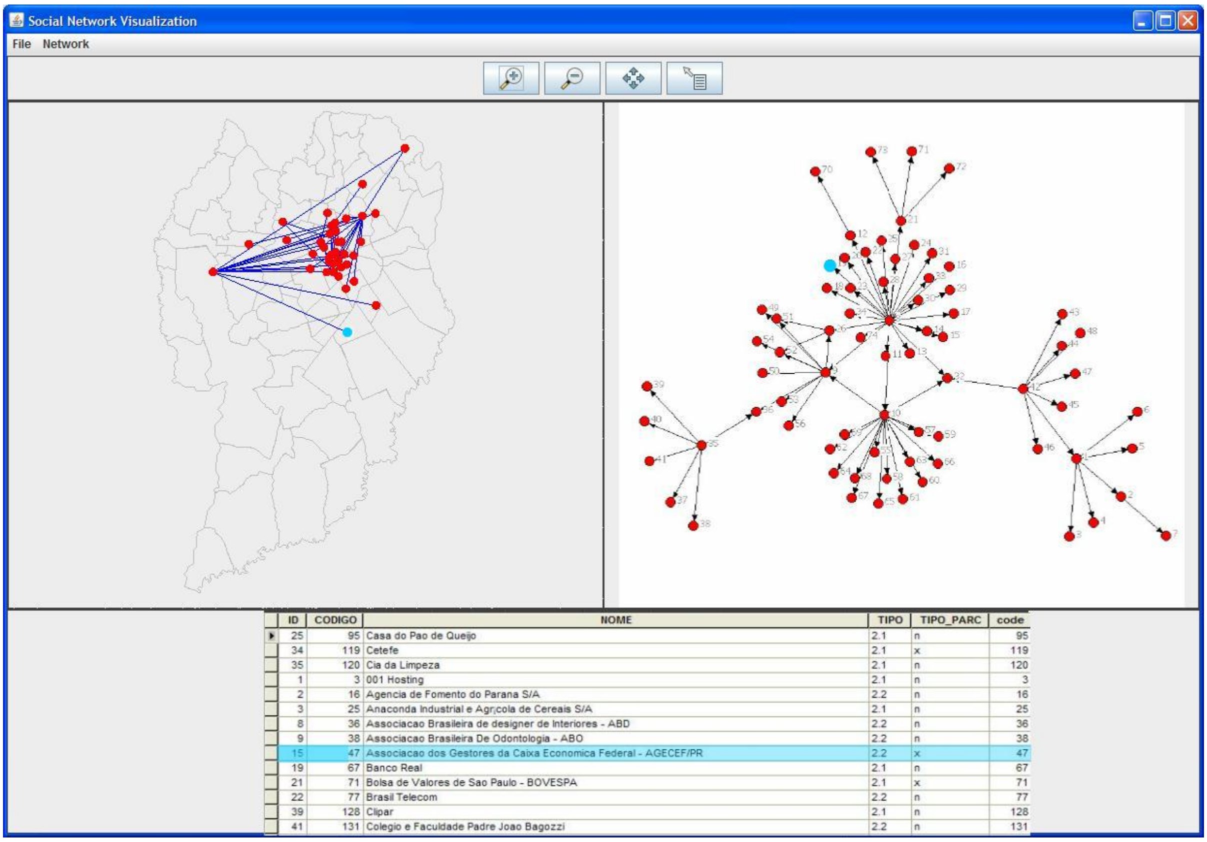


Figure 6 – Different elements connected

5. PRELIMINIARY RESULTS AND FUTURE WORK

The first version of the proposed system is already developed. The basic interface was designed and data regarding social networks, in spatial representation and graphs were inserted. Basic navigation tools (zoom in, zoom out and pan) and query functions were also implemented. The interactivity that allows linking the existing views is being developed.

One of the research' challenges is how to embody cyberspace network representation to the proposed system. This is a difficult task due different levels of detail presented by the data and also because actors in cyberspace may present relationships with other actors located outside the actual geographic reference. So, different scales need to be managed in order to provide this representation. Another issue concerning cyberspace representations regards its dynamicity, which some times is very difficult to follow.

With the use of the system users will realize their analysis more easily considering that all different data representation can be visualized simultaneously. Moreover, new data can be introduced, for instance, census data, in order to allow other analysis.

As a result we believe that a interactivity geovisualization system provides additional insight into underlying information about social networks.

ACKNOWLEDGEMENT

To CNPq for scholarship research **-** Process 308892/2008-9.

Special thanks to Roberto Santos for your help in development of the system.

BIBLIOGRAPHY

ANDRIENKO, N.; ANDRIENKO, G. Coordinated Views for Informed Spatial Decision Making. In: Coordinated and Multiple Views in Exploratory Visualization, International Conference on, 2003. Anais … Zurich, Switzerland. IEEE Computer Society Press. 2003. pp. 44-54.

BRANDALIZE, M. C. B. Metodologia de Mapeamento da Rede de Direito Social Terra e Habitação no Município de Curitiba. VI COLÓQUIO BRASILEIRO EM CIÊNCIAS GEODÉSICAS, 2009, Curitiba. Resumos....Curitiba: Programa de Pós-graduação em Ciências Geodésicas, 2009.

DELAZARI, L. S. Visualização Cartográfica aplicada a análise de redes sociais. 2008. (Relatório de pesquisa)

DELAZARI, L. S.; KAUCHAKJE, S.; PENNA, M. C. (2005). Sistema de Informação Geográfica da Política de Assistência Social do Paraná. In: XXII CONGRESSO BRASILEIRO DE CARTOGRAFIA, 2005, Macaé. Anais..... Macaé: Sociedade Brasileira de Cartografia, 2005.

FREEMAN, L. C. Graphic Techniques for Exploring Social Network Data. In: Models and Methods in Social Network Analysis., P. J. CARRINGTON, J. SCOTT AND S. WASSERMAN. Cambridge University Press, Cambridge, 2004.

FREEMAN, L. C. Visualizing Social Networks. Journal of Social Structure, 1, 1, 2000.

JERN, M. et al. The GAV Toolkit for Multiple Linked Views. In: Coordinated and Multiple Views in Exploratory Visualization, International Conference on, 2007. Anais … Zurich, Switzerland. IEEE Computer Society Press. 2007. pp. 85-97.

KAUCHAKJE, S. Rede sociotécnica de asseguramento de direitos na cidade: proteção social com suporte tecnológico em Curitiba. 2007. (Relatório de pesquisa).

PERER, A. Integrating Statistics And Visualization To Improve Exploratory Social Network Analysis. Dissertation submitted to the Faculty of the Graduate School of the University of Maryland, 2008.

PERER, A.; SHNEIDERMAN, b. Orderly Analysis of Social Visualizations. Social Visualization Workshop at CHI, 2006.

ROBERTS, J. C. State of the Art: Coordinated & Multiple Views in Exploratory Visualization. In: Coordinated and Multiple Views in Exploratory Visualization, International Conference on, 2007. Anais … Zurich, Switzerland. IEEE Computer Society Press. 2007. pp. 61-71.

WASSERMAN, S. et al. Models and Methods in Social Network Analysis. New York: Cambridge University Press, 2005. 344p.