Proceedings of the Fifth IASTED International Conference
VISUALIZATION, IMAGING, AND IMAGE PROCESSING
September 7-9, 2005, Benidorm, Spain

EXPLORING SPATIALLY REFERENCED INFORMATION THROUGH 2D MARCHING GRAPH

Mao Lin Huang and Wu Quan
Faculty of Information Technology
University of Technology, Sydney
NSW 2007, Australia
Email: {maolin, kevingw}@it.uts.edu.au

ABSTRACT

In this paper, we proposed a new visualization framework called Marching Graph that integrates the graph metaphor and the spatial metaphor into a single visualization. Marching Graph allows users to navigate the spatially referenced relational data across two different visual metaphors. We use a force-directed layout algorithm to draw a sequence of progressive graphs, G_1, G_2, ... G_n in a 2D geometric space that present the spatially referenced relational data. Each graph G_i is associated with a particular geographic region R_i presented by the spatial metaphor. We allow the user to "march" through the thematic map by altering the focus region R_i and the display of its corresponding graph $G_i \rightarrow R_i$. The use of 2D visual metaphors facilitates the navigation activities and human cognition process significantly.

KEY WORDS

Information visualization, graph visualization, spatial visualization, and visual metaphor

1. Introduction

With the explosive growth of information available for access, many new knowledge discovery techniques have emerged to meet with the need for the automatic discovery of useful knowledge. Knowledge discovery is a process of discovering interesting and previously unknown, but potentially useful patterns or knowledge from large raw data. The recent estimates show that over 80 percent of available information today contains geospatial referencing (such as geographic coordinates, addresses, and postal codes etc.) [1]. On one hand, various fields of geographic analysis have access to vast digital datasets; on the other hand, numerous conventional datasets also contain either explicit or implicit spatial references. Therefore, many research works [2, 3] have been done in discovering spatial related knowledge from the huge data sets. There are many well-known applications that have been developing in the recent years for, predicting the spread of a disease in NIH, finding crime hot spots in NIJ, and detecting instability in traffic.

Many commonly used spatial related information systems, such as *Geographic Information System (GIS)* provides predefined functions to allow the automatic generation of graphic maps that can be used for viewing, exploration and manipulation of the data sets which contain the spatial attributes. These maps usually employed interactive graphic techniques to enable their interactive capacities. We call these maps the "spatial frame of references". The aim of providing such a graphic map is to amplify the cognition process by displaying a human familiar geometric map as the background to position all spatial related data items on it. Accordingly in the design of input/output user interfaces, most existing spatial data mining methods only consider the use of a single visual *spatial metaphor* (the graphic map) to amplify the cognition process in navigating the geographic locations, while they do not care about the human facts involved in the viewing and exploration of mined data structures (such as a variety of output patterns). In most cases, a simple textual output user interface is provided for users to disclose the underlying patterns. However, in many cases, the outputs patterns are complex, and sometimes are structured with multiple relationships. Much of these data is of multivariate nature, featuring from two to hundreds of attributes. Therefore, a simple textual output cannot satisfy needs to unveil the potential patterns any more. Because of this, there is an urgent need to finding new efficient user interfaces.

Visual approaches to exploration of data analysis have simultaneously arisen in many different areas and now been well applied. The database community has coined the phrase "visual data mining" to describe their efforts, while the terms "exploratory visual analysis" or "exploratory data analysis" are more common within the statistics community. During the above visual exploratory process, information visualization has been used as the vehicle to support the viewing, navigation and manipulation tasks through the visual transformations.

Information Visualization (or Graph visualization) uses a *graph metaphor* to map the extracted patterns into one or more graphs, and then uses the graphic rendering techniques to convert graphs into the visual structure for display. A graph consists of *nodes* and *edges* where the nodes represent the objects of interest, and the edges connect between two nodes according to some association rules. Graph visualization concentrates on the automatic positioning of nodes and edges on the screen. A good layout of graph allows users to understand data more easily. And using an interactive graphics will facilitate users to explore the output patterns and discover underlying information. Existing applications of visual spatial data mining include [4, 5] that combine the information visualization with the thematic maps for the exploration of both the geographic locations and the analytical patterns.

However, most visual spatial data mining methods separate the spatial metaphor and graph metaphor into two visual objects (such as windows, widgets) for display. This, thereby, limits the correlation between two visual metaphors, which will probably cause the following problems:

- **Lack of the overall context of visualization:** if the entire visualization is split into several pieces, then the visualization will lose its completeness. This will greatly reduce the readability and understandability of the overall visual structure. The movements of the focus between different pieces of visualization will cost users extra cognitive effort.
- **Impossibility of cross navigation:** if the entire visualization is split into several pieces, then it is impossible for users to navigate across different visual metaphors which are located in different pieces of visualization.

2. Related Work

To solve the problems mentioned above, M. Kreuseler and H. Schumann have proposed a new visual approach called *Marching Sphere* [6] which integrated the spatial metaphor (spatial frame) and the graph metaphor into one visualization. The graphic visual metaphor is used to visualize the complex structure of mined patterns, while the spatial metaphor is used to display the geographic map as spatial references. In the Marching Sphere, the three-dimensional graphs are generated by an existing tool called KOAN (Kontext Analysator) which can automatically generate 3D graphs by using Virtual Reality Modeling Language (VRML). On the other hand, the Marching Sphere adopts an ordinary two-dimensional interactive map to represent the spatial frame of reference. In order to implement the combination of both embodiments of graphs and spatial frame of reference, Marching Sphere renders the two-dimensional map into a virtual 3D scene. The three-dimensional graph which represents the relational information structure is mapped into the virtual 3D scene and this 3D scene is just located above the area where the interactive map is displayed. Therefore, this achieves the unique mapping between graph-structured information and spatial references.

While the Marching Sphere is a good solution to combining different visual metaphors for visualizing complex data, practically it has the following weaknesses:

- The use of 3D graphs limits its practical applications for ordinary users. Implementation of 3D graphics is computational expensive in ordinary PCs. Therefore, the real-time responses will be slowed down greatly during interactions.
- It is practically difficult to navigate across 2D and 3D visual metaphors.
- The screen and the mouse provided are usually 2D devices. Therefore we won't get the true 3D interactive environment unless we use weird head-gear and expensive bats (flying mice).
- The poor screen resolution available on the ordinary PCs is impossible to render 3D remote objects into sufficient detail for reading; any text in the background is unreadable.

In this paper, we proposed a new visualization framework, called *Marching Graph* that replaces 3D spheres used in [6] with 2D graphs, for the exploration of spatially related data. We use a spring layout algorithm to layout the abstract data structure in a 2D geometric space according to the aesthetics to ensure the quality of layout. At a time we only show the detailed layout of one particular graph above a geospatial area of interest to present the relational data associated with this area. We allow the graph to "march" throughout the whole geospatial map driven by user's interaction. Moreover, we provide a variety of 2D interactive methods (include zooming, filtering, highlighting, coloring etc.) allowing users to interactively navigate through the information space. The use of both 2D visual metaphors amplifies the navigation activities across two kinds of metaphors.

3. A Framework of Marching Graph

The Marching Graph consists of two databases, four processes and two types of interaction (as shown in Figure 1). Our visual interface provides two visual metaphors, the *graph metaphor* and the *visual spatial frame*, allowing users to navigate both the geographic map and the abstract relational structure which references to a particular geographic area that is the current focus on the spatial frame. The spatial frame produced in the visualization provides users with not only the graphic references to the above graph visualization, but also a new progressive spatial-driven navigation method called "*marching*" navigation. It allows users to navigate throughout the whole geographic map and view a sequence of relational structures (graphs) related to the corresponding focused geographic regions by moving the mouse from one sub-area to another progressively.

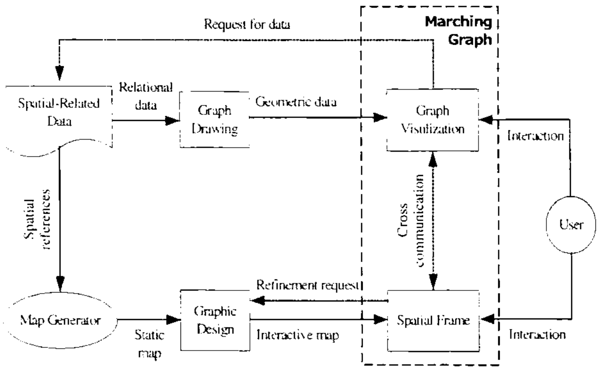

Figure 1: Spatially Referenced Data Visualization Pipeline

We now define these components of our new visualization model:

Data sources:

- *Spatial related data:* contains some spatial references and it could be mined data after the data mining process, and usually contains complex relational structures.
- *Map generator:* provides a detailed static geographic map corresponding to the spatial references extracted from the spatial related data.

Visual Data Processing:

- *Graph drawing:* consists of two steps. First, it extracts the relational structures from the spatial-related data and models these structures into a series of mathematic graphs $G_1, G_2 \ldots G_n$. Each graph G_i is associated with one particular geographical region of the spatial frame. And it consists of a node set V and an edge set E, where the nodes represent the data items and the edges represent the logical relationships among items. Second, it draws these abstract graphs into geometric planes as a series of drawings $D(G_1)$, $D(G_2)$, ..., $D(G_n)$.
- *Graphic design:* extracts essential geographic information from the static map and filters large amount of irrelevant data to form a preliminary thematic map that to be used for spatial frame.
- *Graph visualization:* uses the visual data process method to assign a set of rich graphical attributes, such as shapes, colors, size, thickness, brightness, z-coordinator etc, to the drawings $D(G_1)$, $D(G_2)$, ..., $D(G_n)$ produced in the Graph Drawing phase. It then has to determine the position and display of these graphs.
- *Spatial frame (Interactive maps):* we covert the thematic map generated in the Graphic Design phase into an interactive map (graphics) by using some computer graphic techniques, such as some image design techniques (or tools). We then use the graphical rendering technique to display the interactive map on the screen. This interactive map provides users with a spatial frame for marching through the whole geographic area.

Human Computer Interactions:

- *Interaction with graph:* enables users to visually explore the relational data that might be the outcome of a particular data mining process. The provision of such navigational structures gives users a better understanding of the explicit/implicit relationships underlying the mined spatial-related data.
- *Interaction with spatial map:* enables users to navigate through the whole spatial frame interactively.

Cross Communication between metaphors:

Two visual metaphors are communicating between each other continually during the navigation. At a particular time, there are two corresponding focus areas are activated in marching graph. These focus regions are mutually dependent. Selecting a focus on the map defines the focus of a particular data set, and vice versa. On one hand, the user selects a focused region of interest in the spatial frame, and then a corresponding visual graph is displayed that shows the relational data that has the reference to the sub-region. On the other hand, the user is also allowed to navigate from the visual graph back to the spatial frame.

The details of the two types of interactions will be discussed in section 5. In our implementation so far, we concentrated primarily on the visualization of mined data. As for the use of visualization facilitating the preprocessing and mining process of data mining, we envision it is a further study.

4. Visual Metaphors

Our 2D *Marching Graph* is implemented in Java and it consists of a thematic map (the spatial frame) and a series of circular areas above the map (see Figure 2). At a time only one circular area is fully opened with the green background color and the display of a graph layout $D(G_i)$, while other small circular areas are displayed with the yellow background color as a historic path of navigation (marching) without the display of graph layouts.

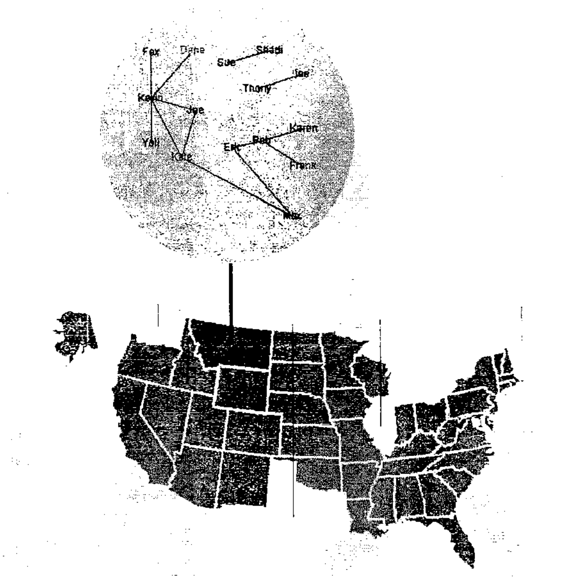

Figure2: A display of Marching Graph with the focus on "Montana" state of USA.

4.1 Spatial Frame (metaphor)

First, our visualization inputs an ordinary geographic source map from the Map Generator (MG) and extracts essential geographic information from the source map and filters the rest of geographic details to create an abstract map. We then use a Hotspot Editor (a free tool) to create a clickable thematic map with reduced size, which can be used to form a spatial frame for navigation.

Particularly we use a Hotspot Editor to define hotspot polygon regions on a bitmap which is based on the abstract map, and save the polygonal data to a file. The hotspot polygonal data can be used for dynamically generating maps at run time. Then we use polygonal data to generate a clickable map. Our system supports powerful poly objects generation. A poly object can have different filling styles, filling colors and border styles. A poly object is a vector object and has many advantages over a raster or bitmap image. It consumes less memory and disk space, and can be scaled to arbitrary size without losing resolution.

4.2 Graph Metaphor

We use graph metaphor to visually present the relational information that are associated with particular geographic regions displayed on the thematic map of the visualization. To generate such a visual metaphor we need to achieve the following steps.

4.2.1 Graph Abstraction

Extract the implicit/explicit relational structures from the spatially related data. The spatially referenced data are very useful which can be used for data analysis in many application domains (such as social economy analysis, health monitoring, etc). These data could be the outcomes of data mining. In many cases, these data is complex and contains relational structures among data items. The modeling of such structures is essential for the analysis and process of data. In this practice, we use graphs G_1, G_2, ... G_n to model these relational structures.

4.2.2 Graph Drawing

To be able to visually represent such relational structures on the screen, we use a physical model called *force-directed layout* algorithm [7, 8] to draw the abstract graphs G_1, G_2,...G_n on a 2D geometric space according to the aesthetics rules that ensure the quality of layouts.

Force-directed layout algorithms use a physical analogy to draw graphs which views a graph as a *virtual physical system of* bodies, where nodes of the graph are bodies of the system. These bodies have forces acting on or between them. Often the forces are physics-based, such as magnetic repulsion or gravitational attraction. Force-directed algorithms aim to compute a locally minimum energy layout of the nodes. This is usually achieved by computing the forces on each node. The forces are applied to each node and the positions are updated accordingly. Force-directed algorithms are often used in graph drawing due to their flexibility, ease of implementation, and the aesthetically pleasant drawings they produce. The experience with force-directed methods shows that they can produce beautiful pictures. They often give highly symmetric drawings, and tend to distribute nodes evenly. This greatly simplifies the cognition process for users to understand the underlying relational structures. Figure 3 shows an example of force-directed layout of graph G_i produced by our visualization.

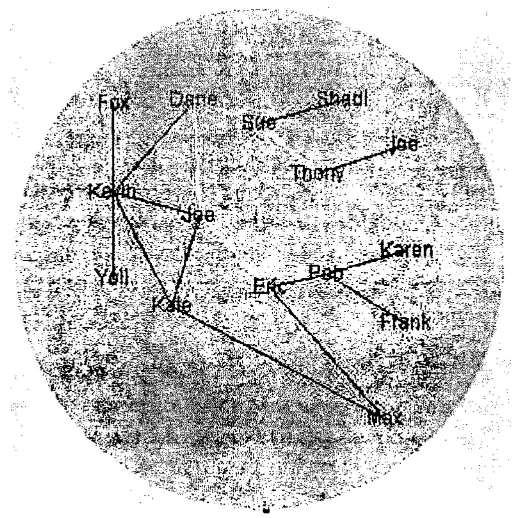

Figure 3: A graph layout produced by Force-Directed Algorithm in Marching Graph

Since each drawing $D(G_i)$ is bounded in a green circular area of a certain size, the display of large graphs with many nodes and edges may cause overlapping problems which will reduce the readability of graphs. To solve this problem, we adopt a dynamic force-directed layout method [9] to maintain the display of a subset of the graph that is currently focused. The dynamic method is used only when the number of nodes excess 30.

4.2.3 Graph Visualization

To display these drawings $D(G_1)$, $D(G_2)$, ..., $D(G_n)$ on the screen, we need to assign a set of graphical attributes, such as *shapes, colors, size, thickness, brightness, z-coordinator* etc, to the drawings for improving of their visibility.

To place the drawing of graph G_i in the visualization, two parameters are needed: 1) the position of its referred geographic region of the map and 2) the amount of display space currently used. The graph position is given by a reference point in the map coordinate system. A graph is positioned by translating its local center point (i.e. the center of a polygon) into its reference point on the map. In order to associate each graph with its corresponding geographic region in the map correctly, we can either position them on the map accordingly (i.e. within the area's borders), or use graphical means such as coloring or connecting lines. We choose the connecting lines to link the graph to its corresponding geographic region. Therefore, when a focus region on the map is selected, the corresponding graph is then popped up above the map with a line connecting to the focus region on the map.

To solve the problem of mutual overlaps among the displays of G_1, G_2, ..., G_n, at a particular time we only display one drawing $D(G_i)$ above the map bounded by a green circular area that corresponds to user's focus. Then as the user moves his focus to other geographic regions, the previous displayed drawing $D(G_i)$ is faded out and substituted with a small symbolic graph (a yellow circle) that is used as part of the history path of navigation.

In order to both keep track of the movement of users' interest and prevent clutters caused by the displays of many small symbolic circles shown at the same time, we limit the number of small symbolic circles displaying on the screen up to five. Then, the visualization allows users to march through the whole map, region by region, for viewing a sequence of graphs G_1, G_2, ...,G_n progressively.

5. Navigation within (and Cross) Visual Metaphors

The success of visual data processing and analysis depends very much on the ability of supporting a variety of exploration tasks provided by the visualization. The visual exploration is the process that users find particular data items through the generation or change of a series of navigational views dynamically and interactively. In our visualization, there are three types of navigation: 1) Intra-Graph Navigation, 2) Inter-Graph Navigation, 3) Marching on the Map (cross metaphors navigation).

5.1 Navigation with Dual-Focuses

In our visualization, two visual metaphors (the graph metaphor and the spatial metaphor) are provided for presentation and navigation of information. The abstract relational information represented by the graph metaphor is located above the clickable map which is represented by the spatial metaphor. Consequently, during the navigation the user usually maintains two focuses. The first focus is on a particular region of the current interest on the map. The second focus is on a particular node (data item) of the corresponding graph that represents the abstract relational information associated with that region of interest.

These dual focuses are independent. The user can march through the map by selecting a sequence of highlighted (or focused) regions on the map. For example, as shown in Figure 2, when a user moves cursor to the state "Montana" on a clickable USA map, this selected region is highlighted with a stronger border and green background color which is different from other regions. At the same time, a detailed graph layout $D(G_i)$ pops up above the focus region showing the relational information related to "Montana" state. We use a thick line to connect between focused region and $D(G_i)$. The user now can move the cursor up to the graph G_i to navigate the relational structure for finding particular data items (relations) of interest. If the user wants to view and retrieve relational information associated with the state "West Virginia", he can do so by moving the cursor over that region. We can see in Figure 4 that the "West Virginia" region is highlighted and the previous highlighted "Montana" region is now semi-highlighted with the yellow color. The previous graph layout $D(G_i)$ is faded away and replaced by a small symbolic yellow circle as part of the history path of the navigation. While a new graph layout $D(G_{i+1})$ is displayed above the "West Virginia" region showing the relational information associated with that region. The "marching on the map" is achieved through the change of a series of focus regions, R_1, R_2, ..., R_i, ..., R_n with the corresponding change of display of a sequence of associated graph layouts, $D(G_1)$, $D(G_2)$, ..., $D(G_i)$, ..., $D(G_n)$.

5.2 Marching on the Map

Our system provides users with a kind of marching effect when users navigate on the map. Particularly, we use the small symbolic circles to show the history of visited graphs. Also the sizes of symbolic graphs are different. The smaller the circle is, the earlier the visiting time is. For example, in Figure 2, the user's focus region is "Montana". When the user changes his/her focus region to "West Virginia", the graph associated with "West Virginia" pops up. Meanwhile, the graph layout displayed above the "Montana" is faded out and is replaced by a small symbolic circle which has the biggest size among all the symbolic circles displayed in the visualization. So, there is a marching effect, as users go through the whole map, region by region.

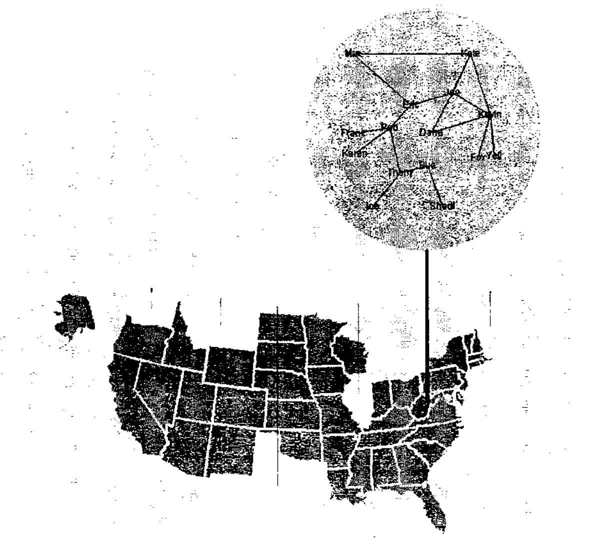

Figure 4: A display of Marching Graph with state "West Virginia" focused.

5.3 Intra-Graph Navigation

Usually the circular display area assigned to a graph layout $D(G_i)$ is limited and unable to display a large number of nodes (say 30+ nodes) comprehensively. Therefore, we adopt "filtering" to provide dynamic display of graphs. Specifically, when a user moves his cursor (or click) on a node, only the direct neighborhood of (or nodes within two hops away from) the focus node will be displayed. Other nodes those are out of this domain will fade out. We use a dynamic force-directed layout algorithm to draw the layout. If the number of nodes is less than 30, we then use the static force-directed layout algorithm to draw the graphs.

5.4 Inter-Graph Navigation

Conversely, we also allow users to navigate across two graph layouts $D(G_i)$ and $D(G_{i+1})$ based on the inter-relationships among graphs. The user can move the cursor over an existing small symbolic circle (as shown in Figure 2) to highlight it with different color ("blue" in our system). If the user clicks on this small circle, then it will be enlarged with the display of detailed graph layout and the corresponding region on the map will also be focused simultaneously.

6. Conclusion

Most visual interfaces separate visual metaphors into different visual containers (e.g. windows) for display. This limits the correlation between visual metaphors and breaks the continuous navigation paths across visual metaphors. In this paper, we proposed a new *Marching Graph* that replaces 3D spheres proposed in [6] with 2D layouts. We use force-directed layout algorithms to draw the relational structure which is much more computational efficient than 3D spheres. We allow users to "march" through the whole geospatial map to find out particular data items by displaying and navigating a series of graphs $G_1, G_2, \ldots, G_n$. Moreover, we provide three types of navigation: 1) Intra-Graph Navigation, 2) Inter-Graph Navigation, 3) Marching on the Map (cross metaphor navigation) allowing users to freely navigate through the complex information space across two visual metaphors.

References:

[1] A. M. MacEachren and M. Kraak, Research Challenges in Geovisualization, *Cartography and Geographic Information Science, vol. 28 (1)*, 2001, 3-12.

[2] K. Koperski, J. Adhikary, & J. Han, Knowledge discovery in spatial databases: Progress and challenges. *Proc. ACM SIGMOD Workshop on Research Issues on Data Mining and Knowledge Discovery.* Montreal, Canada, 1996, 55-70.

[3] S. Shekhar, S. Chawla, S. Ravada, A. Fetterer, X. Liu & C. T. Lu., Spatial databases-accomplishments and research needs. *Transactions on Knowledge and Data Engineering* 11, 1999, 45-55.

[4] G. Fuchs & H. Schumann, Visualizing Abstract Data on Maps, *Proc. Information Visualization, Eighth International Conference on (IV'04)*, London, England, 2004, 139-144.

[5] D. A. Keim, S. C. Noth, C. Panse & J.Schneidewind, Visualizing geographic information: VisualPoints vs Carto-Draw. *Information visualization*, 2003, 2:58-67.

[6] G. L. Andrienko & N. V. Andrienko, Interactive maps for visual data exploration. *International Journal for Geographical Information Science*, 13(4), 1999, 355-374.

[6] M. Kreuseler & H. Schumann, A Flexible Approach for Visual Data Mining, *IEEE Trans. on Visualization and Computer Graphics*, vol. 8, no.1, 2002, 39-51.

[7] P. Eades, A heuristic for graph drawing. *Congr. Numer.*, vol. 42, 1984, 149-160.

[8] T. Fruchterman & E. Reingold, Graph drawing by force-directed placement. *Softw. – Pract. Exp.*, 21(11), 1991, 1129-1164.

[9] M. L. Huang, P. Eades & W. Lai, On-Line Visualization and Navigation of the Global Web Structure, *International Journal of Software Engineering and Knowledge Engineering*, vol. 13(1), 2003, 27-52.