Basak Alper, Selcuk Sumengen, Selim Balcisoy

Dynamic Visualization of Geographic Networks Using Surface Deformations with Constraints

Basak Alper*
University of California at
Santa Barbara, USA
basak.alper@mat.ucsb.edu,

Selcuk Sumengen
Sabanci University, Istanbul,
Turkey
selcuk@su.sabanciuniv.edu

Selim Balcisoy
Sabanci University, Istanbul,
Turkey
balcisoy@sabanciuniv.edu

Abstract: This paper addresses the problem of displaying large network data within geographic context in a 3D interactive environment. Networks often have a spatial component, and maintaining the geographic context of the network data in visualization is imperative for understanding related human behavior and enabling a semantic interpretation.

We propose a novel visualization system, where spatial network data is visualized as deformations on a 3D map. The developed system enables users gleaning dominant tendencies for each node without examining each relation separately. Conventional node and link displays are incorporated for detailed examination of a portion of data on demand. Having two modalities for displaying data as a whole and a detailed portion of it, our visualization system provides a reading for the data at micro and macro levels.

Keywords: information visualization, geographic visualization, 3D deformations, network visualization, force-directed placement

1 Introduction

This paper proposes a visualization system for geographic network data sets, which aims to convey both low level details of the data and high level contextual information with two different visualization modalities.

The first modality, global context visualization represents time-series spatial network data within geographic context through a real-time animation of 3D map deformations.

It maintains spatial framework while providing a qualitative feel of the data by only exhibiting dominant and/or interesting features. The second modality provides a set of interactive analytical tools based on conventional node and link displays, which reveal accurate statistical details of the data on demand.
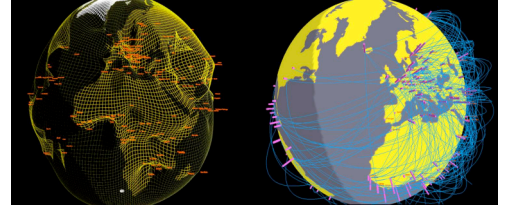

Fig. 1: The diplomatic exchange data visualized with analytical tools (height bars and arcs) on the right versus global context visualization on the left.

The global context visualization technique employs a modified graph drawing algorithm based on spring-embedders which position network nodes according to the time-series data being fed to it. Our contribution lies in projecting complex network data sets into a real-time virtual environment where the geographic framework is preserved. Applied constraints limit variation of network nodes by favoring inherent geographic distribution of nodes. The graph optimization solution is reached using an implicit integration scheme and allows the system to visualize data in real-

time. As the position of network nodes change, the surface is redrawn to fit on the new positions of the nodes. The geographic map projected on the surface deforms and enables viewers to read the data variation as a map deformation. This representation gives a strong qualitative impression and enables viewers to summarize the nature of the data.

Once users have a sense of the overview of the data, they can dig into the details by using analytical tools provided. Interactively responding to users, these tools expand informative quality of the visualization through direct manipulation of visualization parameters. They can explore network data as height bar animations over 3D map or as arcs showing connections. They are able to filter data through selecting nodes or selecting the range of the displayed data.

We examined the proposed method using two different data sets. The first data set is the domestic air flights of US among 231 airports between 1991 and 2004. The second data set is the diplomatic exchange data among 128 nations through years 1815 to 1966.

2 Motivation and Related Work

Visualization of geographic networks requires understanding of the underlying data in its spatial framework rather than the network itself. Graph optimization algorithms or other high dimensional data visualization methods like multi-dimensional scaling do not represent data within a spatial framework, where data nodes are positioned without any geographic constraint [5,12,9].

Conventional geographic network visualization methods superimpose node and link displays on top of a geographic map. [1,2,3] However, for large networks, these visualizations get cluttered with too many line crossings. Averaging, aggregating and thresholding data with adjustable parameterization methods are proposed to reduce clutter. Eick et al. [1] utilized glyphs to encode aggregated data. The disadvantage of losing detailed information about particular nodes was pointed out by Eick et al. and they proposed a matrix representation for detailed data examination which leaves out the geographic framework. The problem of showing both low level details and high level context information is approached by utilizing different visualization modes for each and enabling users to switch between them.

Our visualization system incorporates similar techniques for detailed data examination. However, for global context visualization we propose, instead of encoding data with abstract visual elements, using map deformations for enabling an intuitive interpretation of the data. A similar technique proposed by Dorling et al. [8] makes use of 2D map deformations. Our 3D map deformation relies on a force directed placement algorithm, which animates positions of network nodes according to the time series network data being fed to the visualization. As the position of the data nodes vary, the surface covering these nodes deforms accordingly.

Force directed placement algorithms utilize spring-embedders to solve highly complex graph optimization problems iteratively. First proposed by Eades [6] as heuristic approach, these algorithms seek symmetry, uniform edge lengths, and elimination of long edges in representation of graphs. The resulting graph layout clusters related nodes.

However, maintaining inherent geographic distribution of nodes to some extent is a key factor in our method, which is necessary to guarantee intuitive recognition of the deformed map. Area cartograms representing distribution of data as 2D map deformations incorporate different constraints to reach an optimal compromise for trading shape and area adjustments. The area corresponds to the quantity of the data and some sense of shape must be preserved for recognition [9]. Following a similar logic, we introduced geographic constraints to our algorithm, which limit variation of data nodes, and thus produce a deformed yet recognizable map.

3 Visualization Technique

We propose a geospatial virtual environment for visualizing time-series spatial network data. The first phase of the visualization employs spring-embedders for drawing a graph in which each geographic location corresponds to a single node and non-spatial data components correspond to relations between these nodes. The output graph visualization reflects input data by positioning related nodes closer. In this sense, our technique can be compared to

force-directed placement methods [5]. However, the proposed technique does not follow force-directed placement in any precise sense, but instead exploits its key features. The single most important distinction lies in the geographic constraints applied on the system. These constraints enable the spring-embedder system to reach a configuration that will lead to a deformed map where geographic layout is preserved for assuring intuitive recognition.

In the second phase, the surface covering the nodes is adjusted to fit on the modified positions. As a result, the geographic map projected on the surface deforms and highlights variations in the data. This approach exploits a priori knowledge of the viewer about the physically accurate version of the map. In other words, map deformation facilitates comprehension of non-spatial variables with respect to the geographic framework.

Our approach might be compared to multidimensional scaling (MDS) [11], in the sense that both methods convert high dimensional data relations to Euclidian distance relations on a 3D display. Given both geographical distance relations and data relations (similarity) matrix, our method aims to find an optimal solution for placement of nodes where both geographic context and non-spatial relations are reflected. On the other hand, an MDS algorithm considers only the non-spatial relations and positions nodes to minimize error between input similarity measures and output distance measures.

3.1 Spring-Embedder Model

The spring-embedder model utilized for graph drawing is a modified version of Kamada and Kawai's force-directed placement method [4]. However, in order to maintain geographic framework, we introduced additional connectivity as follows.

In our model, each node is connected to its original geographic position with a zero length spring, the "anchor spring". An anchor spring functions to limit variation of the node and allows a tolerable deviation from its original position. Furthermore, geographically close nodes are connected to each other with uniformly distributed springs, the "neighbor springs", with a heuristic method, where each node is

connected to three closest neighbors with optimally highest angular distinction. The heuristic of neighbor preference is defined by,

$$f(g_i) = \arg\min_{g_j} \left(\sum_k \left(| g_{ik} - g_i | \bullet | g_j - g_i | + C \right) \times \left\| g_j - g_i \right\| \right),$$

where g_i is the current node, g_j is the candidate node and g_{ik} is k^{th} previous neighbor of node g_i. The constant C weighs dot product between vectors (g_{ik} - g_i) and (g_j - g_i) and has a value in range $(1, \infty)$. Larger values of C favor angular distinction over shorter distance. For the current system case C is equal to 2. Neighbor springs aid to preserve geographic contiguity and orientation. Initial lengths of neighbor springs are proportional to the Euclidian distance between the nodes they connect. Therefore, a force applied on a node effects other geographically close nodes, inversely proportional to the Euclidian distance between them. Although this behavior is undesirable in general graph drawing methods, it reflects a meaningful implication for geographic networks. As Tobler [7] puts it, First Law of Geography imposes that "Everything is related to everything else, but closer things are more related." This implies that human activities in a geographic location do not have a discrete effect, but rather have a spreading nature. Springs connecting geographic neighbors reveal this spreading nature.

The anchor and neighbor springs are solely based on geographic data and they exert forces on the nodes to maintain the general shape of the original geographic layout. Besides, each input from the data, which defines a connection between two nodes, is reduced to a three dimensional vector exerting a force on the spring connecting these nodes. These springs are initiated between nodes when a connection is detected in the network data. Their relaxed length is defined proportional to the geographic distance between the nodes. This is similar to Kamada and Kawai's method where relaxed length of a spring is defined proportional to the shortest path between the nodes it connects. As a result, each input from the data deforms the original geographic layout by moving related nodes closer to each other.

We applied geographic constraints, so to speak additional connectivity, instead of calculating repulsive forces as it is the case

in Kamada and Kawai, between every pair of nodes. However, we introduced forces based on relation data and initiated temporary springs between related nodes. Therefore, our algorithms computational complexity is O(V+E), and the number of edges varies depending on the relational data. On the worst case, there might be relations between every pair of nodes and our algorithm has a complexity of O(V²) similar to the Kamada and Kawai's algorithm.

3.2 Force Model

The physical behavior of springs is based on Hooke's law. Therefore, the force exerted by each spring is strictly proportional to the difference of its current and initial length. The simple force equation is as follows:

$$f = k\left(\ \left|n_i - n_j\right| - d_{ij}\right)$$

where n_i and n_j is the current position of nodes and d_{ij} is the geographic distance between them and k is the spring constant.
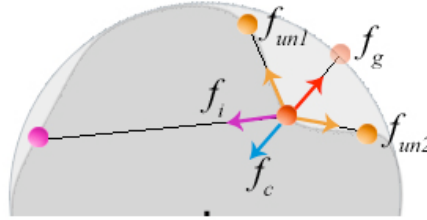

Fig. 2 Simplified force diagram for a node with two neighbors and a single connection data.

Based on the spring connectivity described in above section, there are four sets of forces acting on a single node in each animation step: (i) force exerted by anchor spring f_g (ii) forces exerted by neighbor springs f_{un}, (iii) forces based on connection data f_i and (iv) central force f_c.

$$f(n) = f_g(n) + \sum_{u \in N} f_{un}(n) + \sum_{i \in F} f_c(n) + \sum_{i \in F} f_i(n)$$

Total force acting on a single node n is the summation of above forces for the sets F and N, where F is the set of data inputs in a single animation step, and N is the set of connected neighbors.

The central attraction force is exerted by each input from the data on the corresponding node, and is directed towards the center of the Earth. This force has equal magnitude for each data input and is required to relate the number of total connections of each node with a visual effect, which in this case is the closeness to the center of the Earth. The forces exerted on a single node and their directions are illustrated in Figure 2 for a simplified case with two neighbors and a single connection data.

3.3 Surface Generation

Once the nodes are positioned, the surface covering these nodes is deformed to fit modified positions of the nodes. The surface on which a geographic map is projected is expected to be smooth and fitting all nodes exactly. Initially built as a regular grid, the surface is deformed using a linear interpolation technique based on taking a weighted average of nodal positions within a specified region.

Creating a smooth surface from a high-resolution grid that fits to the graph nodes does not directly lead to a uniform deformed mesh. Therefore, a multi resolution grid approach is introduced: First the surface is built as a low resolution (30x30) regular grid. At this stage, each node is attached to a single grid node that is matched by closest distance relation. Positions of grid nodes that are matched with one or more graph nodes are modified by taking the average of the corresponding network node(s). A discrete linear interpolation method is applied to approximate positions of the grid nodes. Afterwards a higher resolution (120x120) grid is constructed out of the previous surface, and the node matching process is performed again. As the resolution of the grid increases the number of graph nodes that are matched with the same grid node decreases. At the final step, the highest resolution grid (360x360) is constructed. The finer resolution of this grid eliminates matching of multiple network nodes to the same grid node. The interpolation process in this step causes local variations on the surface and makes local fluctuations visible.
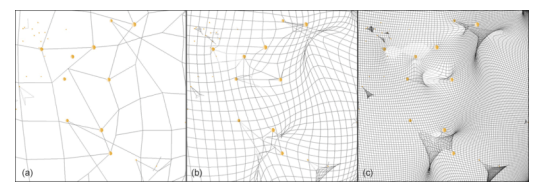
Fig. 3 Multi resolution grid deformation: (a) Low resolution grid constructed with multiple network node attaching to a single grid node, (b) Mid resolution grid reduces multiple attachments, (c) High resolution grid is generated with a Gaussian averaging that covers all the nodes.

Figure 3 displays the steps of the grid formation process. Another contribution of the surface generation method is the topology preservation. Matching of graph nodes with grid nodes are performed using initial coordinates of the graph nodes. This matching will stretch the surface without violating its continuity.
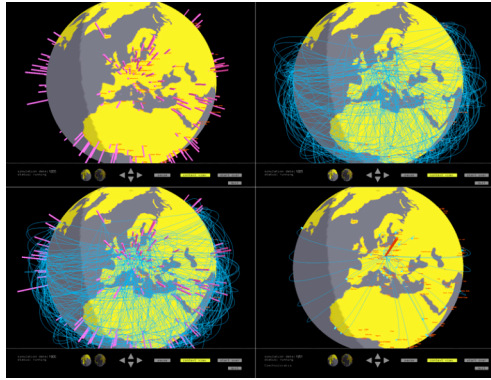
4 Interactive Analytical Tools

Fig. 4 Analytical tools for thorough data analysis. Top left: height bar animations. Top right: arcs connecting nodes. Bottom left: combined view. Bottom right: when a node is selected, only arcs connected to that node are shown. Aggregated load of the node is shown as a height bar.

The visualization environment provides a set of interactive analytical tools, which enable users to drill down into the statistical details of the underlying data. When these tools are active, the deformed map morphs back into its original form. This prevents from occluding the displayed information and decreasing readability.

There are two analytical tools: The height bar animation shows variation in data by animating heights of bars corresponding to each node. If an abrupt change occurs, the height bar of corresponding node is highlighted. When a single node is selected only the height bars for that node and its connections are shown. The second tool renders arcs connecting data nodes. Similarly users can filter arcs by selecting individual nodes. These two tools can be

used together to depict statistics between any two nodes or to analyze network traffic of a single node in detail.

5 Case Studies

The proposed method was analyzed using two different data sets. The first dataset includes records of domestic flights of United States between 1990 and 2004. This data consists of over 80 million flights with ten thousands of different routes among 231 airports. It is freely available from United States Bureau of Transportation Statistics [13].

The output visualization is a real-time, interactive, 3D virtual environment enabling viewers to analyze the flight traffic. The underlying deformation algorithm pushes nodes towards the center of the earth and to each other depending on the flight traffic. The distance between the geographical regions decreases as the in-between flight traffic increases.
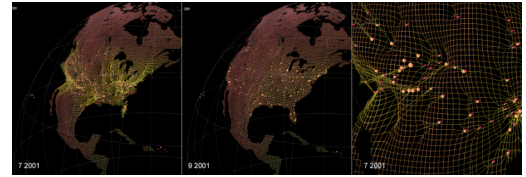

Fig. 5 Snapshots from US domestic air flights data visualization. Image on the left is visualization of July 2001. Middle image displays smoothing of the surface on September 9th 2001. Right image is a closer view for July 2001.

At a glance, the convergence of east and west coasts, especially around southern California in the west and New York in the east, effortlessly conveys information about frequent flights among these locations. The higher curvature of west coast compared to relatively moderate curvature of the east suggests less domestic flights in that region compared to the east.

This method displaying the changes and anomalies over time in the data as a bending, twisting, and relaxing surface, thus utilizes animation as a significant component in the visualization. The effectiveness of animation is best realized when viewing anomalies in the data. For instance, the abrupt expansion of the surface on 9/11 aids the viewer to arrive at a semantic conclusion out of the complex data.

The second data set, the diplomatic exchange data, is obtained from ICPSR database [14]. The data provides information for a total of 128 nations, for the period 1815-1970, the presence or absence of a diplomatic mission from every other member in the set at five-year intervals. The visualization displays political affinities among geographic locations as a physical convergence on the deforming map.
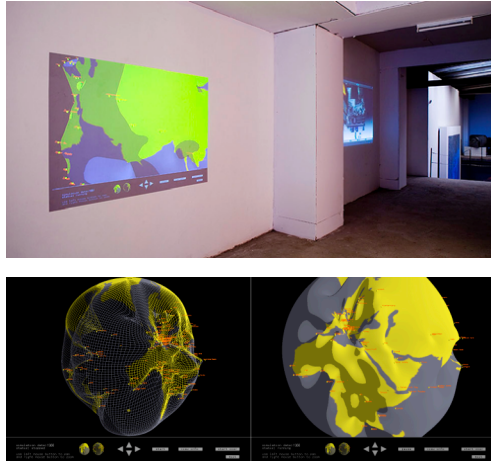

Fig. 6 Visualization system using diplomatic exchange data is exhibited at TECHNE digital performance platform. The bottom two images are snapshots from the application. The top image is the installation where visitors were experiencing the projected 3D visualization using a mouse.

The visualization is exhibited using the diplomatic exchange data in TECHNE digital performance platform for over a week [15]. Users' positive reactions emanate from the visual appeal and intuitive recognition of the world map. Developed interactive navigation tools amplify the engaging quality of the visualization. With the aid of a brief description of the visualized data and the basics of the visualization provided in the help menu, users were able to arrive at simple but valuable semantic conclusions about the nature of the data. The shrink of the surface around Europe, the immobility of unpopulated areas, the decrease of world volume with the increase of political activities are instantly captured qualities of the data. When examined more closely, it is possible to compare political activities of different states. For instance, it is clearly visible that before the 1920's, Canada had much closer political relations with Europe, than US had.

6 Results and Conclusions

Visualization of a large network data in a geographic context requires aggregation and inevitably loss of detailed information when whole data has to be displayed. Our global context visualization is no exception. Yet, substituting abstract visual elements like nodes, links or glyphs with distortion of geographic space has its advantages. The proposed method communicates underlying network data with spatial convergence of data nodes, which is an appropriate physical analogy for visualizing networks. The map deformation emanating from the repositioning of data nodes communicates changes in the data more intuitively and effortlessly. Map deformations are effective because they allow an instant comparison with the mental model of original geographic map.

Moreover, the 3D navigable nature of the visualization system exploits human sensory and cognitive systems at the highest level. In doing so, the visualization does not only depend on the visual cognition abilities of humans, but also utilizes spatio-cognitive skills which are developed to deal with the physical world. In general, navigation is a problematic issue in 3D displays, because users may easily loose a sense of overall context. However, users interacting with a rotating globe are less likely to become disoriented.

The presented system is tested on an Intel Dual Xeon 3.4 GHz workstation with 4GB Memory and NVIDIA QuadroFx 1000 graphics card. The users can navigate in the virtual environment at interactive speeds, 15 fps. to 60 fps., depending on the grid resolution.

7 Future Work

There are several issues requiring further study and elaboration. Improving the poor labeling of geographic locations in global context view will be a significant area of future work. Cartographic labeling methods or focus + context views are possible candidates to develop a more readable and visually satisfying labeling method.

Projecting an opaque or even a semi-transparent map on the deformed surface occludes the position of the nodes and decreases perception. For that reason wire-frame is preferred as the default display

mode since the grid provides the necessary visual cue to the mental model of the map. To overcome occlusion problem and display geographic map simultaneously, interactive widgets are provided to switch between the opaque map and wire-frame mode.

The work presented in this paper addresses public viewers rather than analysts or scientists. Therefore, the focus of the work will remain providing a meaningful big picture of the data to a diversity of viewers. Complex and less intuitive visualizations addressing expert viewers are out of scope of this paper.

In the near future, our goal is to create an installation that works with a live data feed, i.e. transportation or telecommunication data, which benefits public viewers to understand the global context of the data instantly.

References

[1] S.G. Eick, Aspects of Network Visualization. IEEE Computer Graphics and Applications, vol. 16, no. 2, pp. 69 - 72, 1996.

[2] R. Becker, S. G. Eick and A. Wilks, Visualizing Network Data. IEEE Transactions on Visualization and Computer Graphics, vol. 1, no. 1, 1995

[3] K. C. Cox, S. G. Eick and T. He, 3D Geographic Network Displays. SIGMOD Record, vol. 25, no. 4, pp. 50-54, 1996.

[4] T. Kamada and S. Kawai, An Algorithm for Drawing General Undirected Graphs. Inform. Process. Lett. 31, pp. 7–15, 1989.

[5] G. Di Battista, P.D. Eades, R. Tamassia, I.G. Tollis. Graph Drawing: Algorithms for the Visualization of Graphs, Prentice-Hall, Englewood Cliffs, NJ, 1999

[6] P.D. Eades, A heuristic for graph drawing. Congr. Numer. 42, pp.149–160, 1984.

[7] W. R. Tobler, Thirty Five Years of Computer Cartograms. Annals, Assoc. American Geographers, 94(1):58 - 73, 2004

[8] D. Dorling, A. Barford and M. Newman, WORLDMAPPER: the world as you've never seen it before, IEEE Transactions on Visualization and Computer Graphics, Vol 12. No 5., September/October 2006

[9] D. A. Keim, C. Panse and S.C. North, Medial-Axis-Based Cartograms. IEEE Computer Graphics and Applications, vol. 25, no. 3, pp. 60 - 68, 2005.

[10] C. Panse, M. Sips, D. A. Keim and S. C. North, Visualization of Geo-spatial Points Sets via Global Shape Transformation and Local Pixel Placement, IEEE Transactions on Visualization and Computer Graphics, Vol 12. No 5., September/October 2006

[11] A. Morrison, G. Ross and M. Chalmers, Fast Multidimensional Scaling through Sampling, Springs and Interpolation. Information Visualization, vol. 2, no. 1, pp. 68-77, 2003.

[12] C. Chatfield and A. J. Collins, Introduction to Multivariate Analysis. Chapman & Hall, London, 1980.

[13] Bureau of Transportation Statistics, retrieved at June 2005, from: http://www.bts.gov

[14] ICPSR, Inter-University Consortium for Political and Social Research, retrieved at January 2006 from: http://www.icpsr.umich.edu

[15] TECHNE, International Istanbul Digital Perfomance Platform, retrieved at June 2006, from: http://www.techneplatform.org/about.html