SocioScape – Spatio-temporal Visual Analysis of Group Dynamics in Social Networks

BY

MHD KHAIRI REDA
B.S. University of Damascus, Damascus, Syria, 2005

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Computer Science
in the Graduate College of the
University of Illinois at Chicago, 2009

Chicago, Illinois

This thesis is dedicated to my parents.

ACKNOWLEDGEMENTS

I would like to thank Jason Leigh and Andrew Johnson of the Electronic Visualization Laboratory for their support and guidance. Special thanks to Tanya Berger-Wolf for allowing me to start this work and providing valuable suggestions and feedback. Thanks to all the students at the Computational Population Biology Laboratory at the University of Illinois at Chicago for their help and feedback. I would also like to thank Daniel Rubenstein, Siva Sundaresan, and Ilya Fischhoff of the Equid Research and Conservation group at Princeton University, without whom this research would not have been possible. Thanks to Lance Long, Pat Hallihan, and Alan Verlo for a tremendous amount of technical support. Finally, to Luc, Maxine, Dana, Laura, Hyejung, Sungwon, James, Yiwen, Julian, Bob, Ratko, Venkat, Dennis, Ed, thank you for your help, feedback, and support.

MKR

TABLE OF CONTENTS

CHAPTER PAGE

TABLE OF CONTENTS (continued)

CHAPTER PAGE

LIST OF TABLES

LIST OF FIGURES

FIGURE PAGE

LIST OF FIGURES (continued)

FIGURE PAGE

LIST OF FIGURES (continued)

SUMMARY

This thesis presents SocioScape, an interactive visualization tool that embodies a methodology for the visual analysis of spatial and temporal group dynamics in social networks. The methodology introduces a novel visual representation technique suitable for dynamic social networks. This representation provide an advantage over dynamic graphs by explicitly illustrating the evolution of social groups and association choices made by actors over time. The representation is combined with a well-established technique for depicting spatio-temporal data, allowing analysts to investigate the effect of the physical positioning of actors and their movement in the environment on their social behavior. This integration also facilitates the investigation of potential hypotheses that explain the emergence of the observed social structure structure. A case study demonstrates the usefulness of the tool. The primary contributions of this thesis include:

1. A novel visual representation method for dynamic social networks that departs from traditional graph-based visualizations, revealing the evolution of social groups and association choices that actors make over time.
2. A methodology that integrates abstract representation of social interactions with a spatio-temporal visualization to allow analysis of the role of environment in shaping the underlying social structure.
3. A case study in which the methodology was used by behavioral ecologists to explore the social behavior of two animal populations of endangered species.

1. INTRODUCTION

The field of social network analysis has witnessed an unprecedented growth in its applications during the past few years. Its techniques have been used by a variety of scientific fields including epidemiology (8, 9), molecular biology (10, 29), ecology and conservation (5, 7), intelligence and counter-terrorism (23). Social network analysts have also began to incorporate time in their models producing a new class of networks known as dynamic social networks (4).

Visual analysis of imagery depicting social network datasets has been essential to the advancement of the field (15). Although there are a number of well established techniques for visualizing classical static networks, these techniques suffer from limitations when applied to dynamic networks that change with time. Furthermore, these techniques are primarily concerned with visualizing the observed structures the social system being studied. Thus, the vast majority of graphical representations of social networks were purely abstract, and presented out of context of the physical or virtual environment in which the interaction takes place. This makes them ineffective for exploring potential explanations for the emergence of the underlying social structure.

This thesis advances the current state of the art by proposing a novel interactive visualization methodology for the analysis of dynamic social networks in the context of the physical environment in which the interaction takes place. By tightly integrating a spatio-temporal visualization with abstract depictions of the dynamic social interaction, the methodology facilitates hypothesis forming on how the observed social structure emerged in the first place, and how it is effected by external environmental influences.

1.1 Overview of social network analysis

Social networks are abstract representations of relationships between social entities (actors). The most common technique to depict a social network is modeling it as a graph. Actors are represented by nodes, and an edge between two nodes indicate that the two actors are socially linked. The actors in the network could refer to individuals, animals, or institutions, and the links between these actors could be friendship relationships, physical contact, scientific literation co-authorship, or competition. Figure 1.1 shows a synthetic social network visualized as a graph.

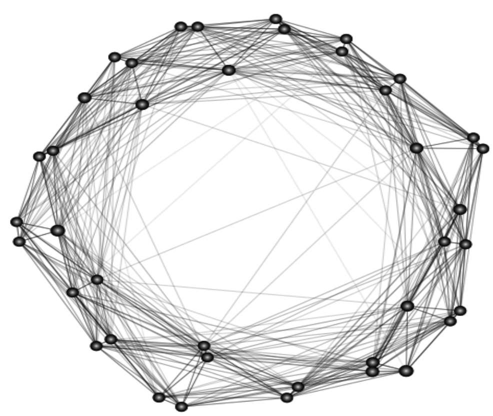

Figure 1.1: Synthetic social network data visualized as a graph. Actors are depicted with circles. A line between two actors represents a relationship between the pair.

Social network analysis is the formal study of social networks by utilizing formal modeling methodologies. A fundamental premise in the field is the understanding that the social structure and the ties between the actors have a major influence on the actors and their behavior within the network (1, 2). Therefore, social network analysts are usually interested in the structural properties of the network, as opposed to the attributes of its actors.

Despite the fact that social interaction is a dynamic process that happens over a period of time, social network models have traditionally treated the network as a static entity. The ordering of interactions between the actors have often been discarded, and the resulting network presented the aggregate social interactions between its actors. That is, a link between two actors is inserted into the network if these two actors were observed interacting at any point in time. This trend to constrains social networks to largely static models was perpetuated primarily by the lack of data collection tools that could accurately sample the social network at fine-grained intervals. Another contributing factor was the lack of sufficient analytical and computational tools to deal with datasets incorporating time.

1.2 Dynamic social networks

Although static models of social networks were, and continue to be, very popular (11), they suffer from a number of limitations when applied to some real-world problems. A static model can not be reliably applied to study dynamic processes such as disease transmission because the model does not address the ordering and concurrency of interactions. These two aspects are very important when studying processes that occur and evolve over a period of time.

Thanks to advances in data collection technologies, social networks can now be sampled at a much higher frequency. This gave rise to a new class of social networks that explicitly incorporated a notion of time (4). Rather than representing the network as a single static graph, a dynamic network can be thought of as a series of graphs, with each graph representing a snapshot of the interactions between actors at a particular moment. Dynamic social networks allowed analysts to study dynamic processes that occur in the network over time. They also offered a new way to think about the social system as an evolving structure.

1.3 Brief history of the development of social network analysis

This section gives a brief history of social network analysis. For a more detailed discussion, the reader is referred to (1) and (2).

The development of the field of social network analysis can be traced to empirical work in sociology and anthropology in the early 18th century. While the analytical methods used in social networks are broad enough to be applied to a wide-variety of fields unrelated to sociology, the development of these methods is strongly grounded in the work of sociologists and anthropologists out of attempts to deal with large amounts of empirical data, and to establish more formal and precise definitions for many of the fuzzy notions that are widely used in sociology (e.g. social groups). Once formal definitions were established to these terms, many of them could be quantified, making it easier to logically reason about the different properties of a society.

The work of Jacob Moreno in the 1930s is widely recognized as one of the foundations of the field. Moreno relied extensively on data collection to study interpersonal relationships. The amount of data he collected soon proved to be incomprehensible without some sort of a formal model providing means to quantitively measure the various properties of social systems. Moreno suggested a set of principles for the development of quantitive analysis methods for social groups, for which he coined the term sociometry. He defined sociometry as "the inquiry into the evolution and organization of groups and the position of individuals within them.".

One of the main contributions of sociometry was the introduction of sociogram. Sociogram is essentially a visualization technique for social networks that is equivalent to graph drawing. It is regarded as one of the earliest examples of graphical representation of social interactions. The invention of sociograms was one of the defining moments in the field of social network analysis. As Moreno puts it "before the advent of sociometry, no one knew what the interpersonal structure of a group 'precisely' looked like" (12). The sociogram model was very simple, yet it was succinct enough to be regarded as a formal model, spurring interest in the systematic formal analysis of social interactions.

Anthropological research also contributed significantly to modern social network analysis. Research by Max Gulckman and James Mitchell on the urbanization and the effect of colonialism of Africa provided additional terminology for describing structural properties of social systems. Terms such connectedness and clusterability were introduced in the 1950s. They remain largely central to social network theory.

In the 1960s, the work of the sociologist Harrison White led to what became known as the "Harvard Revolution" which was credited with establishing one of the fundamental principles in the field: the patterns of association between actors and their position in the network is more important than their individual attributes when predicting the behavior of these individuals. This idea provided a further push towards a network-centric perspective of social systems.

The landmark study by Stanley Milgram in 1967 (26) led to discovery of what came to be known as the small-world phenomenon. This was later found to be a central property in a wide range of networks. The small-world property states that in many types of social networks, a randomly chosen pair of actors are connected by a path that has an average length of 6. Because of the immense practical ramifications of this property, Milgram's work motivated further work on analytical methods and metric for measuring the structural properties of networks. The 1970s saw another widely influential study by Mark Granovetter suggesting that weak relationships are generally more influential than strong ties when it comes to information propagation (59)

Since the 1970s, the field has continued to grow. Many theoretical models have been proposed, along with algorithmic tools for the analysis of networks (2, 11). With increased availability of inexpensive computing resources, the application of these methods have greatly expanded to include areas such as epidemiology (8, 9), molecular biology (10, 29), ecology and conservation (5, 7), intelligence and counter-terrorism (23).

This section has provided an introduction to social network analysis along with a brief history of the field. The rest of the thesis is organized as follows. Chapter 2 provides an thorough review of the techniques used by analysts to produce visual representations of social networks to facilitate exploration and analysis. Chapter 3 introduces SocioScape, an interactive visualization tool that embodies a novel methodology for the analysis of dynamic social networks. Chapter 4 presents an implementation of SocioScape. In chapter 5, SocioScape is used to study group dynamics in two endangered species of wild animals. The effectiveness of the methodology is illustrated with a user study involving ecologists making use of SocioScape to explore the social behavior of these two populations. Chapter 6 concludes with a discussion of the major contributions of this thesis, the new insights gained, and future research directions.

2. STATE OF THE ART OF SOCIAL NETWORK VISUALIZATION

The use of visual illustrations to depict social networks has been a central technique in the field both for understanding the structure of the these networks, and for communicating findings to others (15). It should come as no surprise that the earliest images of social network datasets where based on graphs, with nodes representing actors and edges linking actors who are socially interconnected (12). Despite the substantive body of work on visualizing social networks, the vast majority of these visualization techniques still embrace the use of graphs to depict the social interactions between its actors. While this has proved useful in static networks, experts agree that there are still numerous obstacles to successfully adapting static graph drawing to dynamic networks that change with time (16). Additionally, the wide adoption of social network methods by other disciplines call for new ideas that meet the needs of the new communities of users. Thus, it is becoming essential to design visual analysis tools that combines both abstract depictions of social systems with domain data, creating attribute-rich visualization environments that address the need of domain scientists.

This chapters reviews the state of the art of social network visualization and provides a critique of it. Techniques for static social networks are reviewed first. Follows this is a review of techniques for the visualization of dynamic social networks. This is followed by a review of spatio-temporal visualization methods. The chapter concludes with a discussion of the limitation of graph-based visualizations and makes the case for new depiction methodologies that depart from traditional graph-inspired approaches and integrate domain data into the visualization.

2.1 Static social networks visualization

Static network models have largely dominated the field of social network analysis for decades. Hence, the majority of illustrative visuals are essentially static and provide a purely structural perspective of the network. Consequently, the goal of the visual analysis of static network is two-fold (15):

1. Reveal clusters of strongly linked actors, which are usually referred to as social groups.

2. Reveal the set of actors who play special roles in the network (e.g., one or more prominent actors linking two distinct social groups).

2.1.1 Graph layout

One of the earliest visual illustrations of social networks came from Jacob Moreno (12) who studied friendship relationships among elementary school students. Moreno collected data about friendship preferences of each student, and hand-drawn graphs illustrating these relationships. He also varied the shape of nodes in the graph to encode attributes of the actors. Moreno used his graphs to illustrate differences in the social structure emerging from the interactions of different types of actor.

The first graph illustrations were plotted and drawn by hand, and the position of the nodes was more or less subjectively chosen by the author. With the increased availability of computers, there have been significant effort to develop automatic graph layout algorithms. Graph layout is the process of generating a visual graph from a set of nodes and edges. Layout algorithms generally work by computing node positions and routing edges between them with the aim of producing a comprehensible drawing that allows easier analysis of the data. Layout algorithms differ in the set of underlying principles that are used to judge what a good graph looks like, and the computational techniques used to achieve the layout. There has been a substantial amount of work on graph layout algorithms (13, 24). One of the most commonly used algorithms is force-directed placement (14).

Despite this, the visualization of large social networks is often challenging due to the density of connections typical in many types of networks. Kershenbaum et al (32) identified a number of general approaches that can be used to tackles this problem. They suggest giving the user control over a number of options that affect the visualization including annotation and positioning of the nodes. Although some automatic technique of initial layout is essential, the user should have control and ability to move nodes around. Another useful generic technique is exploiting the hierarchy inherit in most social networks. The user should have the ability to merge highly connected nodes into a single node in order to observe the overall structure of the network. Conversely, a method to unroll these higher level constructs should be made available.

2.1.2 Graph layout toolkits and applications

Graph layout toolkits provides a convenient way of constructing a graph-based visualization pipeline for social networks. They offer more flexibility than standalone visualization applications, but require a certain level of programming experience to use.

Graphvis (22) is one such toolkit. It supports a number of layout algorithms and provide feature to control shape of nodes, their colors, and annotations placed near them. Graphvis reads a graph description file specifying nodes, edges, and annotations in a custom markup language, computes the layout of the graph using the desired layout algorithm, and outputs the graph as a raster or vector image. It is commonly used to generate images of social networks and communicate them to audience. However, its pipeline is designed primarily for offline rendering, making it difficult to be used in interactive applications.

Network analysts started to combine graph drawing algorithms with network analysis tools producing interactive packages that integrated network algorithms, statistics, and graph drawing. These packages were often generic enough to be used in wide variety of fields, including social networks. Pajek (17) and Network Workbench (18) are two such software packages. These packages provide a wealth of network analysis algorithms and statistical methods. However, their visualization features are limited.

Another application designed for the exploration of graphs is GUESS (20). It supports a number of automatic graph layout and statistical tools. It also includes an interactive scripting language that can be used to programatically access the elements (nodes and edges) in the rendered graph to modify their appearance or to add annotations.

The availability of social networking websites such as Facebook and Friendster offered end-users an easy way of building their own virtual social network. Heer and Boyd (17) describe an interactive tool for visualizing friendship relationship of users of social networking websites. The method uses a force-directed, ego-centric graph. An ego-centric layout places the user's node in the center of the graph allowing viewers to focus their attention on their relationships with other actors in the network. To make the visualization more appealing to end-users, special attention is given for aesthetics such as displaying pictures of the individuals represented in the graph and the use of appealing GUI. The application has a number of interactive features including search, filtering, and expansion/collapse of the network.

2.1.3 Graphs in visual analysis of social network datasets

The vast majority of static social network visualizations are done using graphs. A good graph layout with appropriate annotation can be efficient at revealing structural properties of the social network.

A wide spectrum of social networks exhibit what became known as the small world phenomenon. This property was discovered by Milgram (26). Networks that exhibit this property are composed of a number of densely knit clusters of nodes, but at the same time, these clusters are well connected in that the path length between any two randomly chosen nodes is 6 on average. Since this property is common in many networks, graph visualization techniques started taking advantage of it.

Auber et al (25) describe a technique for static graph visualization that exploits the small world phenomenon. Their method relies on a computationally efficient algorithm that filters out edges in a graph, leaving only clusters of highly connected nodes. Each clusters is then collapsed into a single node forming a higher level representation of the network. The resulting cluster nodes are connected together with edges that maintain the topology of the original graph. The technique can be applied recursively until the network is simplified to its simplest form. The user of the system can interactively visualize the network at the desired level using one of the traditional graph layout algorithm. A similar methodology is used in (27). However, the authors employ esthetically pleasing rendering of the graph in 3D space. Their method allows the user to interactively zoom into one of the clusters. As the user zooms in, the nodes spread out revealing the node formation inside the cluster. These two techniques have proven useful for static, small-world graphs.

Force-directed layout algorithms (13) are generally effective at revealing clusters of highly connected nodes, which often correlate with distinct social groups in the network. Suh et al (21) use this property to gain insights into conflicts arising from online collaborative environments. Their work identifies patterns of conflict and mediation arising from article editing on Wikipedia by a large number of editors. A graph is constructed with nodes representing editors and edges between them representing a 'revert' action (a user erasing another user's edits). The authors use their tool to visualize the edit history on some of the controversial articles in Wikipedia. Drawing the graph with a force-directed algorithm shows how the social space is partitioned into distinct opinion groups, including mediators who attempt to reconcile differences among the opinion groups (figure 2.1).

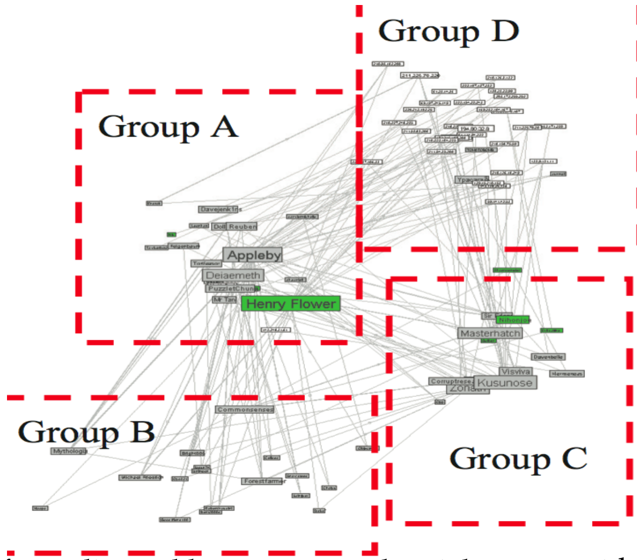

Figure 2.1: Using force-directed layout to reveal social groups with different opinions. Source: (21)

Fisher et al (28) also make use of force-directed layout to reveal patterns of daily collaboration among workers in organizations. Their analysis uses traces of electronic communication to construct a network of individuals who interact together. Visualization reveals recurring patterns of collaboration such as groups of individuals working on a single project forming a highly connected core surrounded by less connected set of consultants.

Heer (31) applies force-directed layout to the employee e-mail communication graph of Enron. His interactive tool allows the user to perform a keyword search and highlights the nodes or edges matching the search criteria. The tool also employs natural language processing techniques to analyze the contents of e-mail messages and classify them into a set of categories (e.g. political influence, company business strategy, etc...). The result of this analysis is used to label the edge representing a particular email with a small pie chart indicating the nature the e-mail's content.

One of the goals of visual analysis of networks is revealing groups of actors who interact closely with each others. However, it is often desirable for the visualization to cluster nodes according to their attributes rather than their connectivity. This is especially useful in domains outside of sociology where experts look for structural patterns among actors sharing a similar set of attributes.

McPherson et al (30) describe an interactive visualization that uses Kohonen's self-organizing memory to spatially cluster nodes with similar feature vectors in an N-dimensional attribute space, placing similar nodes close to each other in the graph (although they are not necessarily connected). The tool also allows attributes of actors to be mapped to the color or size of nodes in the graph.

Although the small-world phenomenon is common in social networks, there are some types of social networks that do not adhere to this property. Another common structural pattern is found in the so called scale-free networks. The node degree in these network follows a power law distributions. That is, nodes with few connections are much more common in the network than nodes with a large number of connections. Typically, these networks do not exhibit a large number of cliques, but rather have a small percentage of prominent nodes that enjoy a large number of connections. Hence, a layout algorithm that facilitates clustering might not produce a good graph here. To deal with this problem, Jia et al (33) employ a method to filter out edges in scale-free graphs while ensuring that nodes that are directly of indirectly connected in the original graph remain connected after the filtering.

The techniques surveyed above provide a way of visually looking at the structural properties of social networks. However, they all assume that the network is static and does not evolve with time. This simply makes them inapplicable to dynamic social networks.

2.2 Dynamic social networks visualization

With the increased availability of dynamic network datasets that incorporate time, there has been a growing need to devise new visualization techniques that allow analysts to see the evolution of networks over time. While the goal of using visualization in static networks was revealing the static social groups within the network, dynamic networks allows us to address the more general question of change over time:

1. How do social group evolve over time?

2. Is there an underlying pattern of social interaction that repeats itself over time?

3. What are the environmental factors that trigger these changes?

2.2.1 Dynamic graph drawing

Dynamic networks can the though of as an extension to the static graph model discussed earlier. A common approach is to think of the network as a series of graphs, with each graph representing a snapshot of the network at a particular moment in time. Each graph have a different set of edges connecting a set of nodes shared by all graphs. Then the problem of visualizing the dynamic network is reduced to the problem of visualizing graphs in which some edges disappear and new edges appear over time. These graphs are known as dynamic graphs.

2.2.1.1 Animating dynamic graphs

A straightforward method to visualize dynamic graphs is animation. Using a graph layout algorithm, each snapshot of the dynamic graph is rendered independently on a separate frame. An important aesthetic principle to consider when animating time-varying graphs is keeping the layout of subsequent frames as close as possible to the first one. That is, the position of the nodes should change as little as possible throughout the animation. If it is essential to move some of the nodes to maintain a comprehensible layout, then the transitioning from the existing layout to a new one should be animated smoothly. This aesthetic element has been referred to as preserving the “mental map”, and has been found important for clear perception of structural changes (24, 47). A variety of techniques have been proposed to achieve a stable graph layout across the animation (34, 48).

Large, time-varying power law graphs pose a challenge to animate because the layout often suffers from excessive clutter, which complicates the effort of maintaining a relatively stable layout throughout the animation. Kumar et al (35) proposes hierarchical stratification of nodes to generate stable clustering of the network. Nodes are organized and rendered in a tree according to their prominence (calculated from the node's degree) with more prominent nodes being at the top of the tree. To alleviate the problem of edge crossing, no edges are drawn between nodes forming a cluster. Instead, nodes belonging to a single cluster are drawn inside color coded bubbles. A unique color is used for each cluster. The technique is effective at visually illustrating the “rise and fall” of prominent nodes and clusters over time. However, less evident is evolution of these clusters (i.e., how nodes interact inside a cluster, and how it moves from one cluster to another over time).

2.2.1.2 Temporal unrolling of dynamic graphs

A different approach that avoids animation is unrolling the evolution of the graph over time in the spatial dimension (38). For each timestep, a graph is drawn using a conventional layout algorithm in a plane. The planes are then stacked on top of each others to illustrate evolution of the network, with latest timesteps being at the top. The planes are drawn semi-transparently so that planes of earlier timesteps are still visible. The method is helpful in pointing out differences between two consecutive timesteps. However, this technique is limited to visualizing few timesteps at once. It becomes hard to "see through" changes going further in time as the stacking of slices increases opacity. Figure 2.2 illustrates this technique.

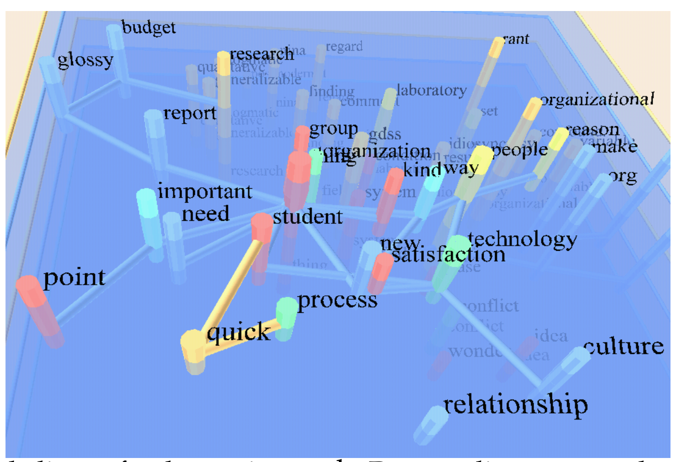

Figure 2.2: Stacked slices of a dynamic graph. Recent slices are rendered on top of older ones. Source: (38)

Gaertler et al (39) use a similar method, but avoid drawing the 'white space' plane on which the graph is grounded. This allows for more slices to be stacked while maintaining visibility of earlier slices. On the other hand, this leads to difficulty in perceiving the ordering of nodes (i.e., which nodes and edges belong to which timesteps). It is also susceptible to cluttering.

2.2.2 Analysis of group evolution

Static graphs are usually effective at showing groups of actors who interact closely with each others. Although the same techniques could be used to highlight social groups in individual timesteps in dynamic graphs, it is hard to see how these social groups evolve. As time passes, new groups emerge and older groups dissolve. Additionally, a number of groups might merge forming a larger group or a group could split into smaller groups. These changes can be evident in an animation if they happen over a short period of time. But what if those changes happen over long periods? An observer will likely miss such structural changes if they are spread over few minutes in the animation. Moreover, with graph animation, it is difficult to spot actors switching their membership from one social group to another simply because of the sheer number of nodes. On the other hand, temporally unrolling the dynamic graph allows one to see momentarily changes happening over few timesteps. However, simply pilling up more timesteps degrades the visibility, making the technique ineffective for analyzing structural changes that happen over longer periods of time. To address these two issues, a number of techniques have been proposed to allow analysis of the evolution of social groups, and to track actors' affiliation with these groups.

2.2.2.1 Interactive analysis of dynamic interaction graphs

Yang et al (36) discuss an interactive tool for visualization of dynamic graphs coupled with an automatic event detection frameworks. The framework identifies structural changes involving social group such as two groups merging together or one group splitting into two, which can then be highlighted in the visualization. The user can select a node and visualize its connections using a series of subgraphs representing the node's neighborhood over time. A limitation of this technique is that although it can detect timesteps at which certain interesting events occur, there is no easy way of tracking changes to groups in terms of actors membership over time since the framework does not provide a notion of a stable tag for the detected groups. That is, two groups at consecutive timesteps sharing a similar set of actors are treated as two completely independent groups.

2.2.2.2 Analysis of subgraph evolution

Falkowski et al (37) proposes a model that tracks evolution of groups by employing cohesion and stability statistical measures and plotting there curves over time. For networks in which actors' group membership is fluctuating, a visual method is used to tag similar groups and track their evolution over time. The technique uses a 2D graph layout in which the X axis represents time, and the Y axis is used to arrange the different groups. Groups are depicted using nodes with edges linking two or more groups merging or splitting. This depiction is useful in illustrating patterns of group evolution in the network. However, there is no way of finding out the role of individuals in this evolution, as actors are not depicted in this visualization. Additionally, there is no way of tracking the membership choices actors make over time (i.e., which groups a particular actor chooses to associate with and how this association changes).

2.2.2.3 C-Group

Kang et al (40) describe C-Group, an interactive tool for studying group association for a pair of actors over time. The user selects two actors from the network which are referred to as the focal pair. The visualization highlights the shared and non-shared individuals with which the focal pair interact. These associates are grouped together according to their attributes. The user can also specify the time window form which the associates are displayed. The advantage of this technique lies in the clean and effective layout (figure 2.3). A limitation of this approach is although it allows one to see shared collaborators between two authors, for instance, and how these collaborator change over time, it does not show the relationship between those collaborators. This makes it ineffective for exploring social groups and their evolution over time. Another downside is that the group semantic is defined in terms of two focal actors only, and it can not be easily extended to include a notion of a larger communities of actors closely interacting together.

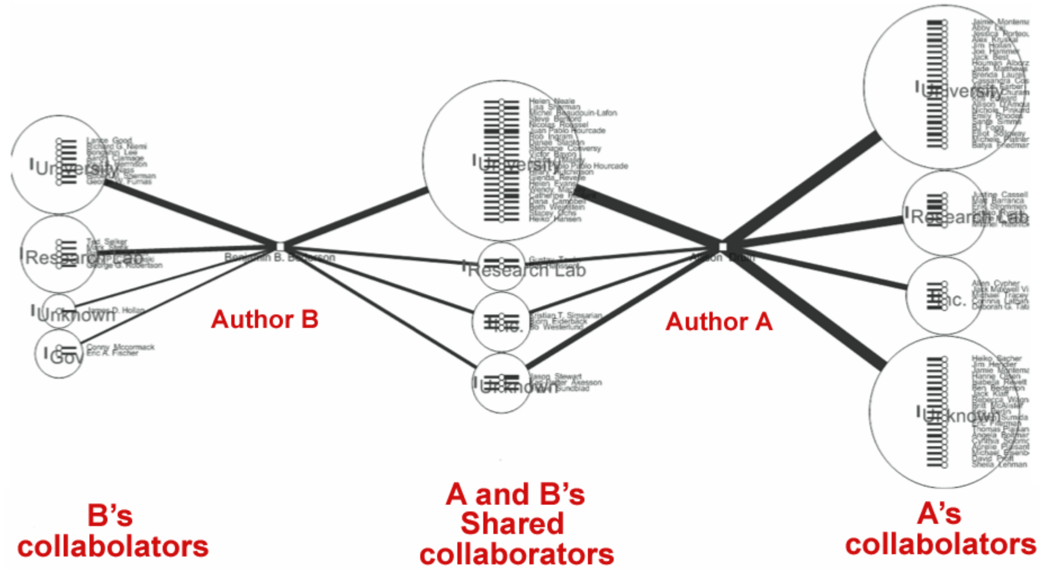

Figure 2.3: Screenshot of C-Group showing co-authorship relationships of the focal pair of authors. Collaborators who coauthor papers with both A and B are displayed in the middle, grouped by their institutions. Authors who collaborate with only one of the two focal authors are displayed on the side. Source: (40)

2.2.2.4 Community interpretation graph

Tantipathananandh et al (3) use a special kind of graph to depict the composition of social groups at different timesteps. Every actor is depicted with a unique node at each timestep. The nodes are organized into layers with each layer representing one timestep. The layers are drawn on top of each others with layers earlier in time drawn near the top of the diagram. Each node has exactly one incoming and one outgoing edge that link it to its two siblings representing the same actor in the previous and next timesteps, respectively. This allows tracing an actor's group affiliation over time in the graph by following these edges. Actors belonging to the same social group are surrounded by solid, color-coded boxes. Figure 2.4 illustrates this technique. Unfortunately, this technique does not adhere to the principle of preserving the mental map. That is, the position of actors in the visualization could change even though the actor remains associated with its original group (e.g., actor 0).

Additionally, the technique is susceptible to cluttering due to edge crossing, which makes it difficult to trace the position of actor across all timesteps.

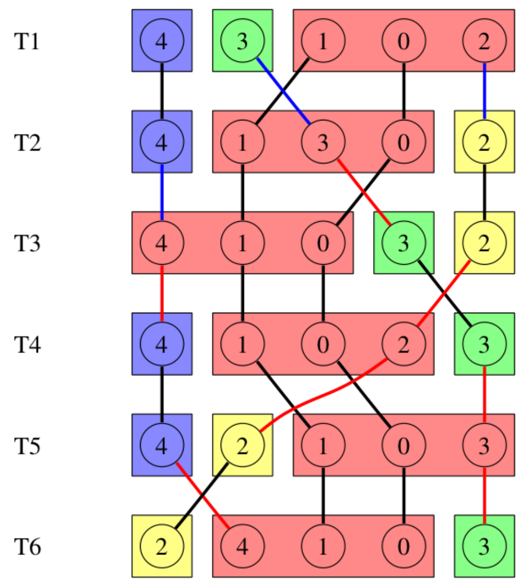

Figure 2.4: A graph-variant showing the composition of social groups over time. Groups are depicted with colored rectangles encompassing actors belonging to those groups. Colored edges depict an actor switching association from one group to another. For example, actor number 3 leaves the green group to join the red group at T2. Source: (3)

2.2.3 Hybrid approaches

Although graph drawing remains the predominant way of visualizing the dynamics of interaction in social systems, there has been some work on augmenting graphs with other visual representations to make is easier for analysts to recognize trends and patterns that occur over time.

Viégas et al (41) describe an interactive tool for visualization of e-mail interaction aimed at end-users. Their system combines traditional graph layout of the user's contacts network with a 2D diagram illustrating the temporal rhythms of e-mail interaction between the user and each of her contacts. Animation is employed to show how the patterns of interaction evolves over time. As time passes, contacts move up in the 2D space indicating periods of intense conversations, or descent toward the bottom indicating a slowing of e-mail exchange.

In a separate project, Viégas et al (42) introduce a technique for analyzing the dynamics of collaboration on Wikis. The system comprises a tool that tracks revisions to Wikipedia pages along with a visual depiction method similar to parallel coordinate diagrams. Revisions to a particular article are depicted using parallel columns arranged on the X axis according to their date. The columns are further divided into segments stacked on top of each other with each segment depicting a portion of the article. Identical segments that survive revisions are highlighted by drawing horizontal lines between columns. This illustration shows patterns of collaboration between authors such as "edit wars" (marked by a zig-zag pattern).

Although these techniques have shown promise in revealing some of the patterns that occur in social spaces, they are largely application specific, and can not be easily extended to other disciplines. Therefore, it is difficult to draw from them a general set of guidelines on the type of depiction methodologies that can be effective at revealing patterns of interaction among social actors in a more generic social setting.

2.3 Visualization of spatio-temporal data

A large portion of the data that exists today is geographically or spatially tagged one way or another. Some researchers suggest that up to 80% of data that exists in corporate databases is geographically referenced (50). Many classes of social interactions are shaped by the spatial displacement of its actors with in the environment. Example of social networks that are affected by the geographic location of actors include disease transmission networks and scientific collaborations. Therefore, taking the spatial attributes of actors into account while studying their interaction over time might reveal the role of their physical location in shaping the dynamics of interaction. In fact, revealing the spatio-temporal patterns of interaction in the network is one of the primary goals of domain scientists who use social network models.

Visualization is probably the most effective and natural way for the communication and exploration of geo-spatial data. This is because our visual processing system has evolved primarily to allow us to quickly analyze the various objects that occupy the space surrounding us. In the last few decades, the amount of geo-temporal information that we are able to collect and record has increased exponentially. Therefore, there is an increasing interest in interactive visual analytic tools that can be used to show the data spatially as well as how this data changes over time. These techniques have been referred to as spatio-temporal or geo-temporal visual analytics.

There is a large body of work on the visualization of spatio-temporal data. However, there is no research yet that attempts to combine dynamic social network visualizations with spatio-temporal analysis techniques. If done effectively, the combined analysis of the social interaction in both the temporal and spatial dimensions could reveal many interesting aspect on the role of geography in shaping the social structure and its evolution over time. This section will review some of the work on geo-temporal visualization techniques.

A straightforward way to visualize the changing position of objects within space is using animation (52). However, the disadvantage of using motion to visualize change stems from the fact that the visual working memory of humans is very short term. Therefore, these techniques face the same challenges that also hinder the use of animation in dynamic graphs.

In 1970, Hägersrand (51) proposed the notion of space-time cube to visualize events the occur in different places at different times. In this technique, the base of the cube is used to represent the spatial dimension, while the temporal dimension is represented in the 3rd dimension of the cube (usually the vertical one). At the time this work was published, its real-world applications were severely limited simply because it was time consuming to produce such illustrations by hand or print. However, technical advances in computer graphics led to the resurfacing of this method. Soon enough, interactive, computer-driven visualizations started using the space-time cube to visualize spatio-temporal data (53, 55). Andrienko et al (54) propose using the space-time cube to visualize the movement of entities over time. Some of the recent applications include intelligence, surveillance, and other military applications (56, 57).

One of the main advantages of the space-time cube is its ability to collectively display multiple events or objects simultaneously in the same visualization space, and track the position of these objects over time. Changes to the objects' attributes over time can also be easily illustrated with color-coding, for instance.

2.4 Critical analysis of graphs

Graph-based visualizations have proven useful for revealing some of the structural properties of static social networks, including social groups and prominent actors. The most widely used layout algorithms are spring and force-directed techniques. These algorithms work naturally to produce a layout that shows clusters of highly interconnected nodes, which are usually correlated with social groups. Another advantage of using graphs is the availability of well developed and efficient algorithms that are able to cluster nodes on multiple levels, producing a hierarchical view of the network. Combining this with an interactive visualization pipeline that adheres to the zoom, filter, and detail-on-demand paradigm allows easy exploration of the various social groups in the network.

Despite being so common, graphs have a number of limitations that often hinder their usefulness when used to visualize large datasets. Moreover, latest trends in social network analysis characterized a shift towards dynamic models and datasets further increases the complexity of graph-based representations. These limitations are discussed next.

2.4.1 Reduced visibility due to clutter

All graph drawing methodologies suffer from cluttering one way or another. As the number of nodes and edges in the graph increases, the number of edge crossing increases at a faster rate. Empirical evidence suggested that minimizing edge crossing have a significant positive impact on the readability of a graph (45). Thus, increased edge crossing often lead to difficulty in answering basic question about the structural properties of the network such as whether two nodes are connected (directly or indirectly). Additionally, displaying annotations for nodes or edges further increases occlusion and degrades the overall quality of the image.

One solution that can be used to reduce this problem is to offer options to the user to reduce or increase the mount of displayed nodes or annotations. This can be coupled with clustering algorithms that can produce multi-level hierarchical views of the network (32). However, these techniques have been applied mainly to static graphs, and it is not clear if they can be easily extended to dynamic graphs that change with time.

2.4.2 Effect of graph layout on drawn conclusions

The positioning of nodes in the graph has been found to have a significant effect on structural properties inferred by the viewer. Blythe et al (44) have shown that the perceived importance of a node in the network is negatively correlated with its distance from the center of the layout. Similar factors have also been found to affect the number of distinct social group recognized by the viewer. This suggests that using two different graph layout algorithms to visualize the same social network might lead to different conclusions being drawn about the network. Moreover, there is no simple criteria for selecting the algorithm best suited to visualize a particular network.

On the other hand, most of the graph drawing algorithms are designed to geometrically optimize the layout according to a set of aesthetics, such as minimizing edge crossing and maximizing symmetry, which are usually determined by the algorithm designer. These principles, however, are most of time based on the subjective reasoning of the algorithm designer, and usually are evaluated according to their computational complexity. Very little work has been done on how effective these principles are to produce diagrams that enhance understanding network, or to critically analyze them from the point of view of the visual cognitive system. Moreover, combining different aesthetic principles does not necessarily correlate with increased or decreased readability of the graph (46).

2.4.3 Computational complexity

Producing a high quality graph layout is computationally expensive. For example, force-directed layout algorithms have a complexity of $O(n^3)$ where n is the number of nodes in the graph (24). This hinders the use of these algorithms to render large graphs. It also places limitations on their usability in interactive visualizations.

2.4.4 Tuning of graph layout algorithms

Graph layout algorithms require a relatively large number of parameters to be set by the user. Although most end-user tools provide a way of automatically setting those parameters, or provides a preselected set of them, those values are intended to work with a wide variety of network, and will not necessarily produce an optimal layout for every network (16). Therefore, the user is faced with the option of accepting these parameters and having to put up with a less-than-optimal layout, or directly adjusting these parameters. In most cases the users of the system are not experts in graph drawing, and the affect of the different parameters on the layout is usually not well understood by most users.

2.4.5 Preserving the "mental map"

One of the important aesthetic principles in dynamic graphs is preserving the "mental map". This refers to maintaining the position of nodes throughout the animation with as little change as possible. Empirical evidence have shown that adhering this principle when rendering dynamic graphs have a positive impact on readability (47). However, it has been recognized that this principle often conflicts with other common aesthetic principles such as minimizing edge crossing. This degrades the quality of the layout in individual time slices as the mental map preservation principle puts constraints on moving nodes, leading to increased cluttering. A number of solutions have been suggested in an attempt to compromise between these conflicting principle (48, 49). However, the problem manifests itself with larger graphs, and in cases where node and edge annotation is desired.

2.4.6 Ability to reveal patterns of interaction over time

There have been no work on the effectiveness of animation in dynamic graphs in revealing patterns of social group evolution, or actor group association. However, evidence from studies on the human visual perception can be extrapolated to point out potential limitation of the use of motion to depict changing relational information. The human visual working memory is typically short. It is difficult for the viewer to remember the structural changes that are unfolding as the animation plays. Thus, many of the patterns governing structural change that happen over time are likely to be missed as the viewer will not be able to relate previously witnessed changes to the unfolding ones. Moreover, The viewer's attention will likely be focused on a very few number of interesting structures on the graph. Therefore, if the dynamic graph contains a moderate number of interesting structures, the user will be forced to pick a few of these structures and focus her attention on them throughout the animation. If the viewer decides to inspect other structures, the animation will have to be replayed, increasing the effort and time required to study the whole graph.

2.5 Summary

Although there is a wealth of research on visualization of social networks, the majority of approaches still embrace graphs as the primary depiction method. While there are some advantages to using graphs including the availability of a wide variety of graph visualization algorithms and tools, there are still a number of challenges that need to be overcome before graphs can be effectively used to depict dynamic social networks. The approach proposed by this thesis relies on novel visual representations that are designed from the bottom up to address the temporal aspect of social interaction in real-world environments. The methodology also takes into account the requirements of the broad community of domain scientists that routinely make use of social network analysis in their research. Table 1 compares the techniques surveyed in this chapter against the proposed methodology.

TABLE I

COMPARISON OF THE SURVEYED VISUALIZATION TECHNIQUES AGAINST THE METHODOLOGY PROPOSED IN THIS THESIS

Approach	Includes time	Reveals evolution of groups over time	Reveals actors' group membership choices over time	Spatio-temporal analysis	Filtering	Includes statistics
Static graphs	-	-	-	-	Yes	Yes
Animated dynamic graphs	Yes	Limited	-	-	Limited	-
Unrolled dynamic graphs (38, 39)	Yes	Limited	Limited	-	Yes	Limited
Subgraphs evolution (37)	Yes	Yes	-	-	-	Yes
C-Group (40)	Yes	-	Yes	-	-	-
Dynamic interaction graphs (36)	Yes	-	Yes	-	Yes	-
Community interpretation graph (3)	Yes	Limited	Limited	-	-	-
This thesis	Yes	Yes	Yes	Yes	Yes	-

A rethinking of traditional graph-centric methodologies is not only useful, but also imperative for the effective depiction of dynamic social networks (43). Furthermore, the wide adoption of social network analysis by domain scientists makes it important to design visualization environments that strongly couple domain specific data with abstract social interaction depictions in order to show how the social behavior of actors and groups is affected by the physical or virtual environment in which the interaction takes place.

3. METHODOLOGY

This chapter proposes SocioScape, an interactive visualization tool that embodies a novel methodology for the temporal and spatial analysis of group dynamics in social networks. First, a set of requirements are established to help guide the design of the tool. After this, the visualization pipeline is discussed. Finally a detailed description of the visual representation techniques used in SocioScape is presented along with a discussion of their advantages over graphs.

3.1 Analytical goals of SocioScape

In dynamic social networks, analysts seek to uncover patterns that govern the evolution of social groups and the behavior of actors over time. Although sociologists are usually interested in the structural outcome of this evolution, other domain scientists are concerned with understanding the environmental factors that influence the behavior of actors, and contribute to shaping the structure of the network. For example, ecologists need to understand how the social behavior of animals is affected by their reproductive state, resource requirements, or the presence of human activity within their habitat. Epidemiologists might be interested in the effect of a local contamination on the spread of disease in a particular location. Public watch groups seek to understand how campaign contributions influence voting of elected officials. All of these are example of social systems in which the behavior of actors is, more or less, affected by their environment, and by other external factors.

The vast majority of visualization techniques for social networks assume a static model in which the network does not change. These techniques are therefore inapplicable to dynamic networks that incorporate time. Although some techniques exist for visualizing dynamic networks, these techniques focus almost exclusively on revealing the social structure present in the network. Little attention is paid to the context in which the interaction occurred in. Thus, these techniques are often ineffective for explaining how and why the observed structure emerged in the first place. For example, although there are some tools that can visualize the social structure inherent in wild population of animals, these tools do not attempt to combine environmental attributes with the visualization (such as the position of individual animals or the location of water holes). Therefore, an ecologists can not use these tools to explore the role of resource distribution, among other factors, on the social behavior these populations.

There is a need for a new visualization methodology for dynamic social network that go beyond graph-based representations and address the needs of domain scientists. In order for the methodology to be effective, it should meet a number of key requirements. First, the methodology should explicitly include some notion of change over time. Its graphical representations should depict the evolution of social groups, as well as association choice actors make (i.e., what groups an actor chooses to associates himself with over time). These representations should also be easy to understand and interpret, even for users who are not experts with graph-based representations. The attributes, along with their implications, are discussed in detail in the rest of this section.

3.1.1 Dynamic

Social behavior is a dynamic phenomenon that evolves with time. Therefore, any visualization methodology that attempt to be useful for network analysts should acknowledge this fact, and explicitly provide semantics and visual depictions that allow the representation of dynamic interaction between its actors.

Another requirement stems from the diversity of social systems that can be modeled using dynamic networks. Social interactions in these systems can occur over varying time scales ranging from minutes to years. Therefore, the framework's definition of social interaction and time should be flexible enough to capture interactions at both fine and coarse-grained temporal scales.

3.1.2 Depiction of social groups evolution

A fundamental phenomenon that social network analysts look for is social groups. These groups are composed of a number of actors who interact closely with each other. Examples include research groups in an academic institution or a group of friends sharing a particular hobby. In dynamic social networks these groups evolve over time. As time passes, new actors join a social group, and existing actors leave the group to join another one.

Understanding the dynamic evolution of these groups is imperative to understanding the overall structure of the society being studied, and how this structure evolves over time. Hence, the framework should provide effective visual depictions that facilitate analysis of group evolution.

One of the challenges in visualizing the evolution of groups in dynamic networks is that the very notion of these groups has traditionally been static and transient. That is, groups appearing in different timesteps are considered more or less independent. Without an appropriate dynamic group semantics, it is hard to come up with useful visualizations that illustrate the development of these groups. Hence, the framework should not simply visualize the transient groups that appear at different timesteps, but should introduce a persistent notion of a social group that is more flexible in terms of actor membership. We make use of the model proposed by Tantipathananandh et al (3). This model established the notion of "community", a grouping of individuals that persist over time, allowing new members to join in and exiting members to leave. It also provides a stable tagging scheme that can label instances of the same group at different timesteps. The stable labeling of communities allows generating a stable layout that adheres to the principle of preserving the mental map.

3.1.3 Depiction of actors' group membership choices

Another related analytical goal is studying association choices that actors make over time. In social systems, actors continuously make decisions about whom to interact with, what social groups to join, or whether to leave their current group. These association choices when taken collectively form the fabric of the social structure, and influence its evolution over time.

An understanding of the patterns that govern individual behavior in the society takes the analysts a long way towards explaining why certain actors have more or less influence on the network, and how the micro-interactions in the society give rise to its structural elements. Therefore, the visualization framework should provide mechanism to effectively depict the association and group membership choices that actors make over time.

3.1.4 Integration of spatial and temporal analysis

In many examples of social networks, the geographic placement of actors within the environment have a significant impact on the form of social interaction that takes place. This fact has often been neglected by most social network visualization tools as most network analysts were traditionally interested only in the sociological aspects of the interaction with not much regard to the physical environment. However, as domain scientists start to rely on these tools, the need to explicitly address the influence of external environmental factors on the social behavior is becoming more important.

An effective depiction of the spatial arrangement of actors and groups in the environment is imperative to domain scientists who are not only interested in the emerging social structure, but also demand an explanation that takes environmental factors into account. The visualization framework should also be easily extendable to include other domain attributes relating to actors and groups.

3.1.5 Easily comprehensible depictions

Although graphs are powerful at abstracting many relational information, they are not easily comprehensible by non-experts. As more domain scientists start to use social network models, there is a growing need to come up with more user friendly visual representations that are easier to understand and analyze.

This section discussed a set of key attributes that the proposed methodology should meet and provided justifications for these attributes. The next section describes a visualization pipeline designed to meet these attributes.

3.2 Visualization pipeline

A visualization pipeline for social networks describes a set of transformations on the raw social interaction data to produce a graphical representation of that interaction. The proposed pipeline adheres to the Haber-McNabb visualization model (58). Figure 3.1 shows the a schematic description of SocioScape's visualization pipeline. First, the data is preprocessed using a community identification algorithm to detect the social groups in the network. The result of this step is referred to as the community interpretation. The number of communities present in the community interpretation can be overwhelming. To reduce the amount of data that ends up in the final visualization, filters can be interactively applied to select a subset of the detected communities, or to limit the analysis to a short time-window. Finally, the data coming out of the filters is rendered using two visualizations – the space-time cube which depicts the physical movements of communities, and the Affiliation Timeline which depicts the evolution of communities over time. The elements of the pipeline are described next in detail.

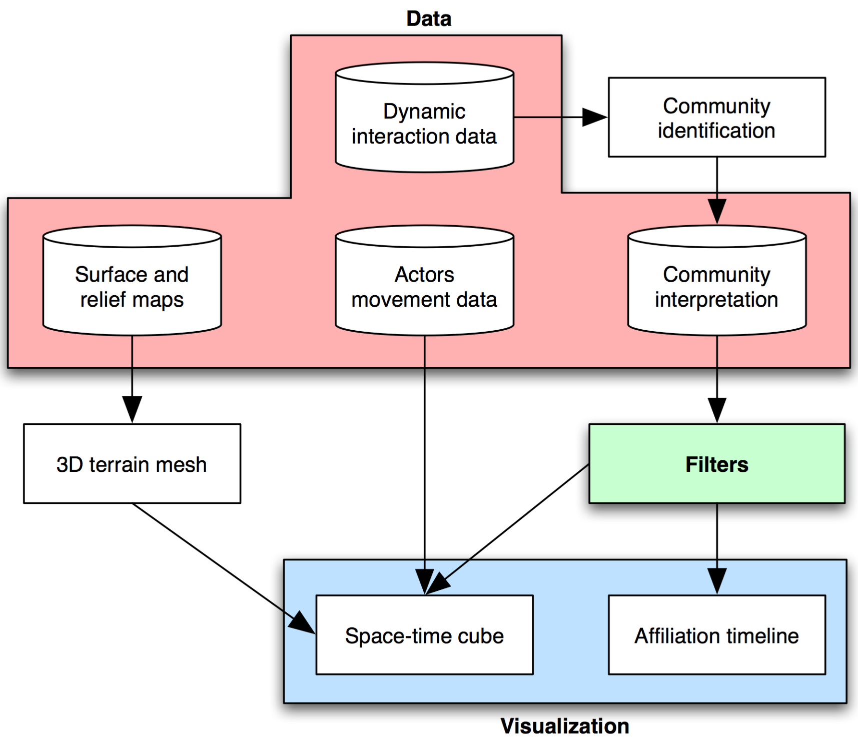

Figure 3.1: SocioScape's visualization pipeline

3.2.1 Dynamic interaction data

The dynamic interaction data is a digital representation of a real-world social network. The social network is observed for a continuous period of time. The network is sampled an arbitrary number of times during this period generating a static snapshot at each of these timesteps. The sampling process does not have to be periodic. This means that the time period between two consecutive samples could vary.

Each static snapshot is represented by a static graph specifying cliques (complete subgraphs) of actors that were observed interacting with each other at that snapshot. This means that the unit of interaction in our model is a group of actors transitively interacting with each other, as opposed to interactions between a pair of actors. This constraint is imposed by the community identification framework (described in the next section) that SocioScape uses. Although this interaction model is more restrictive that the traditional dyadic relationship model, it nevertheless captures a wide range of real-world social interactions. On the other hand, the benefit of this restriction from a visual representation perspective that it enables us to represent the overall group dynamics, as opposed to addressing the micro-interactions that occur inside a group.

3.2.2 Community identification

This step implements the community identification framework proposed by Tantipathananandh et al (3). The algorithm detects the unique communities present in the network after taking into account the collective interactions between all actors over the observation period. A community is grouping of actors that persists and evolves over time, allowing new actors to join and existing ones to leave the community throughout the observation period. Each community is given a unique tag (or color). At each timestep, a community is represented by one of the transient groups of actors who were observed interacting with each other at that timestep. However, the notion of community is temporally stable. That is, a community usually spans more than one timestep though its members are likely to change over time. Figure 3.2 presents an example outcome of the community identification process.

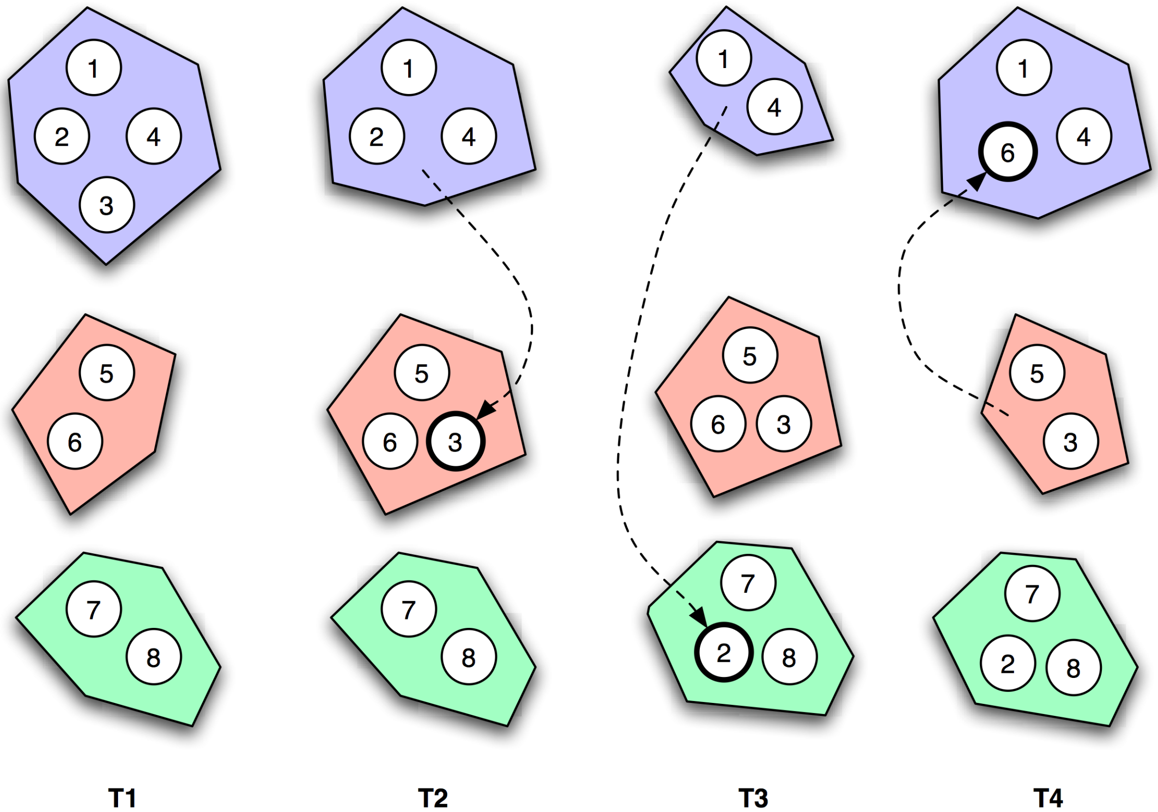

Figure 3.2: An example output generated by the community identification algorithm described in (3). At each timestep (T1 through T4), actors who were observed interacting with each others are grouped together. Three unique communities are detected (blue, red, and green). At each timestep, every one of the three communities is represented by exactly one group. The definition of community allows its members to switch their affiliation from one community to another over time. However, an actor can be affiliated with one community only at a given timestep. Affiliation switch is indicated by an arrow from the previous community to the new one.

3.2.3 Community interpretation

The community identification process is carried out as a preprocessing step as the algorithm described in (3) is computationally expensive. The result of the algorithm tags each of the groups at each timestep with the a community ID (unique across one timestep). This data is referred to as "community interpretation".

3.2.4 Filters

There are three types of filters that can be used by the user to limit the amount of information that ends up in the visualizations.

- Temporal filter: This filter allows the user to specify a continuous time-window that falls anywhere within the observation period. The length of the time-window can also be specified by the user, and can be expanded to cover the entire observation period. Only communities and actors that were active during this time-window will end up in the visualization. This allows the analyst to focus her attention on a small time period for detailed analysis. Conversely, one can visualize the interactions in the entire dataset at once and look for an over all pattern.

- Actors filter: This filter allows the user to select a subset of actors to be visualized. Actors that were not selected do not show up in the visualization.

- Communities filter: Similar to the actors filter, the communities filter allow the user to select a subset of the detected communities to be visualized.

3.2.5 Affiliation Timeline

The Affiliation Timeline is one of two visualizations modes in SocioScape. It comprises a 2D visualization that is similar to a parallel coordinate diagram. The visualization depicts the evolution of communities, as well as the association choices that actors make over time. The temporal dimension is depicted on the X axis instead of using animation, while the communities and individuals are arranged in the Y axis. The depiction methodology used in the Affiliation Timeline is discussed in detail in section 3.3.1.

3.2.6 Actors/group movement data

Often the sampling process in dynamic social networks records not only the social interactions between actors, but also their geographic location within the environment. If this information is available, it can be used in the space-time cube visualization showing how actors and groups move during observation period. The movement data is composed of a database that stores the physical position of each of the communities at each timestep using some coordinate system.

3.2.7 Surface and relief maps

High resolution geographic maps are becoming increasingly available as the number of satellites providing commercial mapping services increases. It is easier than ever to obtain surface and relief maps (also known as Digital Elevation Models) for the environment in which the social interaction takes place. This data can be used to enrich the space-time cube visualization, allowing domain scientists to analyze the effect of geographic positioning of actors and groups along with other environmental influences on the social behavior.

3.2.8 3D terrain mesh

Using a relief map, a 3D terrain model can be generated. The model can then be used in the space-time cube visualization allowing domain scientists to examine the effect of topography on the movement and interactions between groups and actors.

3.2.9 Space-time cube

The space-time cube is a 3D visualization based on the work of (51, 53). It shows the movement of communities over time. The space-time cube is discussed in detail in section 3.3.2.

3.3 Visual depiction of spatio-temporal group dynamics

SocioScape uses two representations to visualize the spatio-temporal group dynamics in the subject social network. The first representation, referred to the Affiliation Timeline, depicts the evolution of communities, and illustrates how actors switch their affiliation with communities over time. The second depiction is based on the space-time cube (51, 53). It depicts the movement of communities in the environment. Together, these two techniques are combined in an interactive visualization environment enabling a network analyst to analyze the patterns and trends of social interaction over time. Additionally, the spatio-temporal depiction of movement in the space-time cube allows domain scientists to answer questions related the effect of the geographical positioning of actors within the environment on their behavior.

3.3.1 Affiliation Timeline

The Affiliation Timeline is a 2D representation that resembles a parallel coordinate diagram though with different semantics. The visualization depicts two phenomena of interest:

1. The evolution of communities in the network.

2. The association choices that actors make over time. That is, the communities that a particular individual chooses to associate himself with at different timesteps.

Figure 3.3 shows an instance of the Affiliation Timeline depicting the communities illustrated in figure 3.2. The X axis represents time, while communities and actors are arranged on the Y axis. Communities are depicted with non-overlapping rectangular areas (each rectangles represents exactly one community), and the different rectangles are arranged on top of each other. These rectangles are divided into “slots” which are occupied by actors affiliated with the community. Actors are depicted with lines that fall within the community with which they are affiliated. At a particular moment in time, all actors falling within the same rectangle (i.e., the same community) are said to be affiliated with that community at that timestep. This representation groups actors based on their shared community affiliation, as opposed to drawing edges between them to indicate mutual association. This helps avoid the cluttering caused by overlapping edges in graphs.

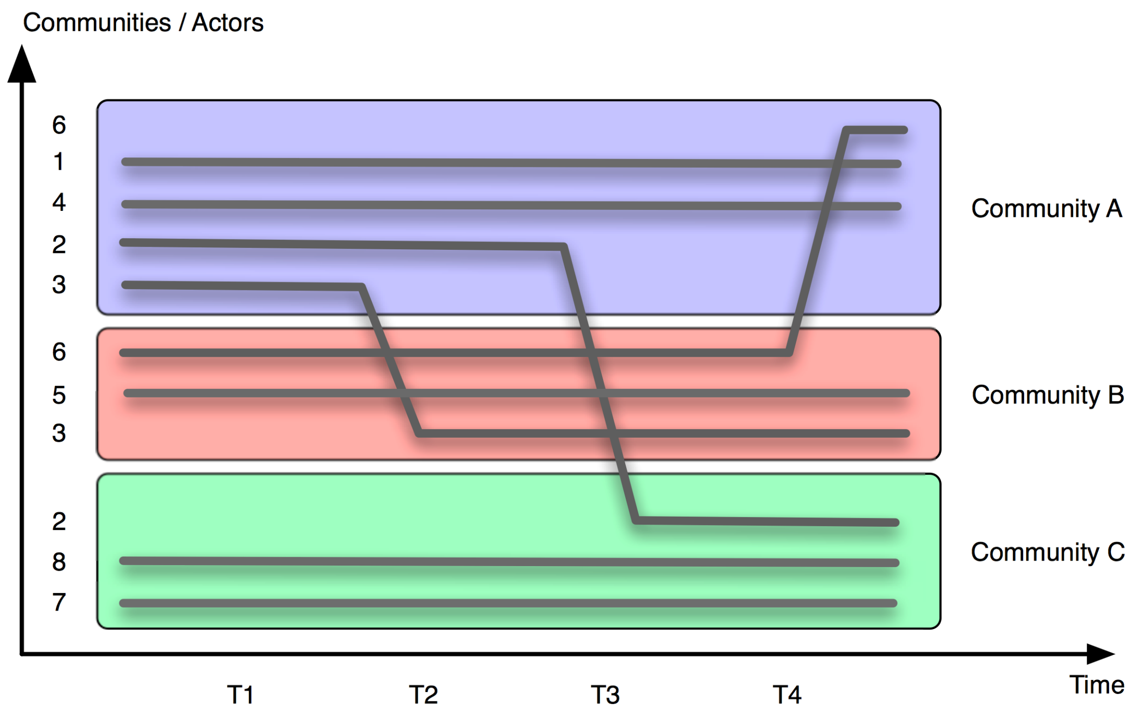

Figure 3.3: Affiliation Timeline depicting communities in the example presented in figure 3.2. The X axis represents time, while communities and actors are arranged on the Y axis. Actors are depicted with contiguous lines that fall within the community with which they are affiliated. Labels on the left associate lines with actors' ID, while labels on the right indicate the different communities. A skewed line indicate an actor switching its affiliation from one community to another. For example, actor number 3 leaves community A and joins community B at timestep T2.

Each actor has exactly one contiguous line. A straight horizontal line indicates that an actor remains affiliated with its community. When an actor leaves its community and joins another one, its line is skewed towards the new community at the timestep at which the switch occurred. If an actor leaves its community at one point in time and later returns to it, the actor returns to the same slot it had originally occupied. At the moment, the vertical arrangement of actors and communities inside the diagram is based on a first-in-first-out principle. However, the layout could be later optimized to minimize line crossings.

There are a number of factors that make the Affiliation Timeline efficient at depicting group dynamics. These factors are discussed next. To visually illustrate these factors and prove the Affiliation Timeline's superiority over dynamic graphs, we shall compare figure 3.3 to figure 3.4, which depicts the same dynamic social network data illustrated in figure 3.2 using a graph layout instead.

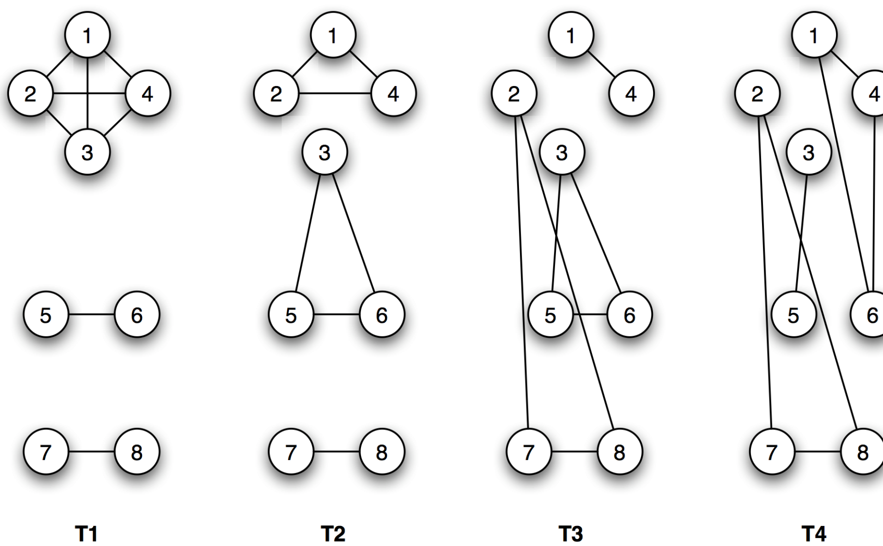

Figure 3.4: A graph layout of the dataset depicted in figure 3.2

3.3.1.1 <u>Perceivability of social groups</u>

In the Affiliation Timeline, communities are clearly marked in the visualization by enclosing them with non-overlapping rectangles, and the different communities are color-coded with unique colors to help differentiate them. The human visual processing system is very efficient at segmenting enclosed, non-overlapping geometrical structures. This should make it easy for the analyst to examine the contents of these rectangles at any point in time. Hence, the task of determining the actors affiliated with a particular community should become perceptually intuitive at all timesteps.

In contrast, recognizing the different communities in the network is not an intuitive task when the analyst is presented with an animated graph layout that preserves the mental map across timesteps. As the community structure in the graph changes with time, the relative positioning of nodes no longer reflects those communities (i.e., nodes closer to each others might not necessarily belong to the same community). For example, in figure 3.4, determining the different communities in timestep T4 is not intuitive. Rather, it requires the analyst to sequentially trace the edges connecting nodes to each others in order to see if they belong to the same community.

3.3.1.2 Perceivability of actors association choices

Another common task for network analysts is determining the different communities that a particular actor choose to associate himself with over time. The Affiliation Timeline could again provide an intuitive mechanism for tracing an actor's association history.

By unrolling the temporal dimension and depicting each actor as one, solid, contiguous line, the user can easily follow actors as they move between different communities over time. This is because the visual processing system of the viewer quickly segments the different lines that belong to different actors in the early stages of the cognition pipeline, allowing the viewer to quickly trace an actor's community affiliation history. When an actor switches affiliation form one community to another, the straight line suddenly becomes crooked, attracting the attention of the analyst.

In contrast, spotting an actor switching affiliation from one community to another is not an easy task when the network is depicted using a graph. For example, in figure 3.4, it is not easy to notice that actor number 6 switched affiliation and started interacting with actors 1 and 4. In figure 3.3, this phenomenon is far more evident as the line representing actor 6 suddenly jumps from community B to community A.

Another feature that the Affiliation Timeline provides is the ability to see the entire affiliation history of an actor throughout the observation period (or a portion of it). However, this is very difficult to see in a graph animation as the visual working memory of humans is very short. On the other hand, temporally unrolling the entire graph is feasible only for a few timesteps. Otherwise, the graph suffers from excessive cluttering.

3.3.1.3 Filtering and layout stability

SocioScape provides robust temporal filtering semantics allowing the analyst to select a contiguous time-window of an arbitrary length that falls anywhere within the observation period. This updates the Affiliation Timeline to render only the affiliation history of actors within the specified time-window. The space occupied by each timestep on the X axis is expanded to cover the available space in the diagram. Figure 3.5 shows the same Affiliation Timeline but with a time-window spanning T3 – T4.

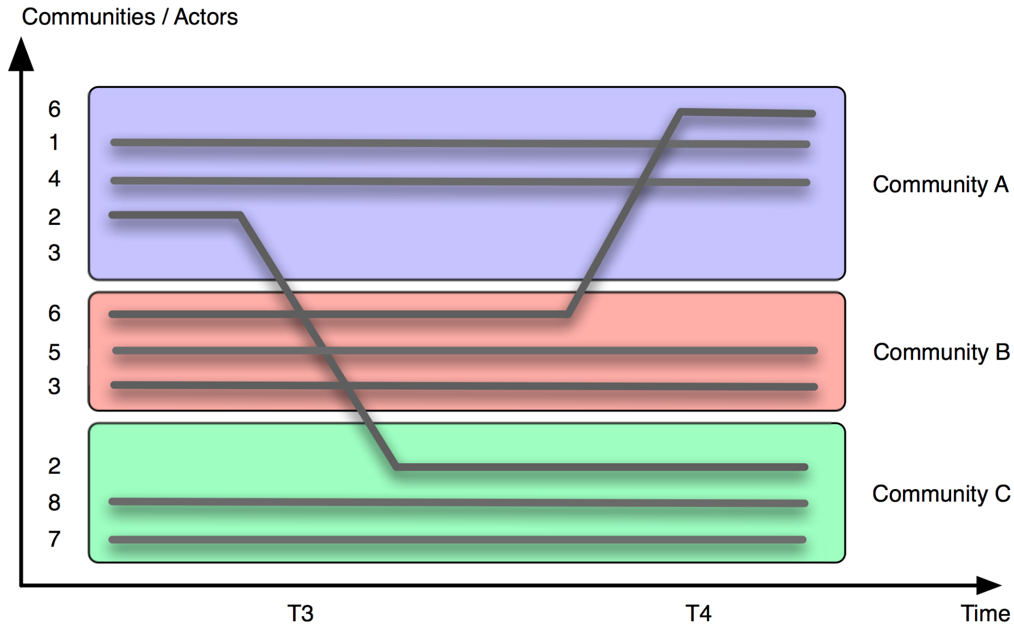

Figure 3.5: Affiliation Timeline depicting the same data shown in figure 3.3 but with the time-window spanning T3 – T4

An important property of the Affiliation Timeline is that the vertical position of communities and actors within the diagram is invariant with respect to the selected time-window filter. That is, a when the user chooses to shrink the temporal filter to focus on a shorter period of time, the position of actors and communities does not change. This is because the vertical position of actors and communities is first pre-computed to cover the entire observation period, creating all the necessary slots inside each community to hold actors.

The stability of the layout throughout all timesteps allows the user to interactively explore the social interaction at different scales simply by adjusting the length of the filtering time-window or its position within the observation period. With a fully stable layout, the user's mental map will be preserved, minimizing distractions and increasing the effectiveness of the visualization.

3.3.2 Space-time cube

The space-time cube visualization is a 3D representation that is based on the model proposed by Hägersrand (51) and Karrak (53). It depicts the movement of communities in the physical environment over time. Figure 3.6 illustrates the space-time cube visualization. The base of the cube (XY plane) depicts the spatial dimension, while the Z axis is used to depict time. The cube's base is a 3D topological map of the region. An abstract map, or a satellite image can be rendered on the 3D map if the data is available.

Communities are depicted using spheres. The horizontal placement of communities (in the XY plane) reflects the physical position of the community in the environment at the time of sighting, whereas its vertical position reflects the time of sighting. Recent sightings and positioned higher than older one. A line connects two consecutive sightings to depict the movement of the community. A vertical line is projected from the spheres onto the topological map to disambiguate the actual position of the community sighting within the environment.

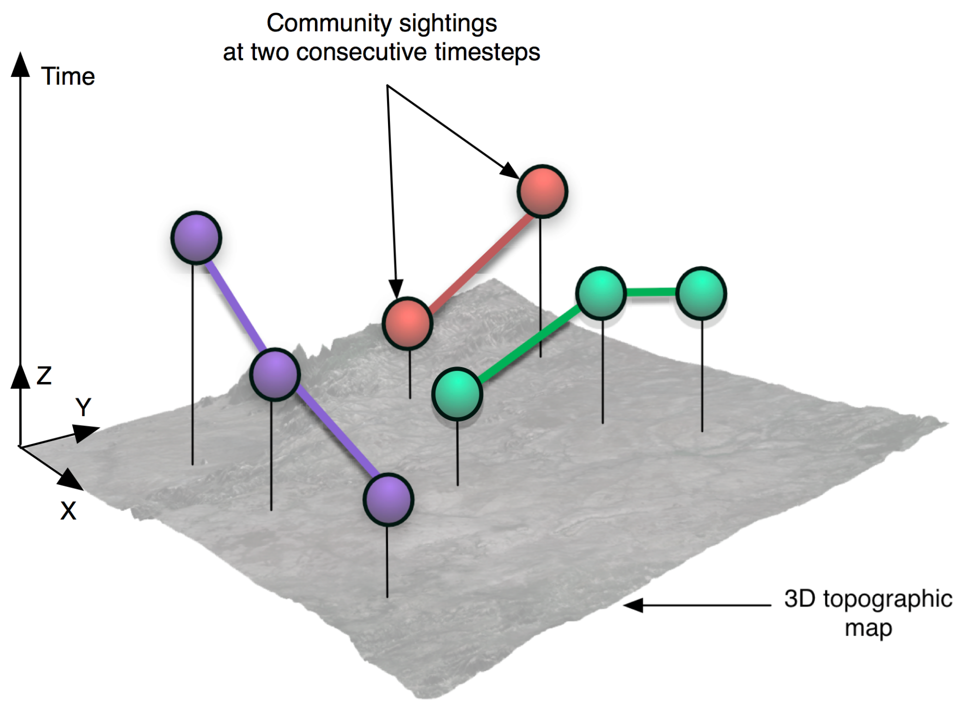

Figure 3.6: Space-time cube visualization

The space-time view can be interactively manipulated by the user. The vantage point from which the scene is rendered can be translated in the 3D space, or by zooming in or out. Additionally, the whole scene can be rotated in 3D space. This allows users to change perspective in order to disambiguate some of the inferences. It also allows the user to focus on some of the sightings by zooming into the scene, or see an overview of the movement by zooming out. The main contribution of the 3D space-time cube to the visualization is that it shows how the different social groups share the environment and physically interact with each others over time.

Often, the social interaction is significantly affected by environmental factors such as the geographical position of groups and their relative distances to each others. Combining the space-time cube with the Affiliation Timeline provides a method for exploring both the social behavior, and some of the environmental factors under which this behavior takes place. This is helpful for domain scientists who seek to uncover how the environment shapes the structure of the society being studied, and how that structure responds to environmental changes over time. For example, it allows an ecologist to see how herds of animals compete to access resources in a shared habitat, what strategy these groups employ to avoid predation, or how the behavior of these groups adapts to changing ecological factors such as temperature.

3.3.3 Combining spatial and temporal analysis of communities

SocioScape combines the Affiliation Timeline with the space-time cube in an interactive visualization environment. The two visualizations can be used to display the same social network side-by-side, allowing an analyst to see both the community structure, as well as the geographical position and movement of these communities over time.

To amplify the user's ability to analyze both the spatial and temporal aspects of the social interaction, SocioScape provides a semantic cross-highlighting feature. Data selected from one of the diagrams causes related data to be automatically highlighted in the other. For example, the user can select one of the community sighting locations in the space-time cube visualization by simply clicking the sphere. This in turn causes a portion of that community's rectangle to be automatically highlighted in the Affiliation Timeline. This highlighted portion shows the actors affiliated with that community at the time of sighting. Conversely, the user can highlight a period in the Affiliation Timeline by drawing a box.

This causes the community locations sighted during that period to be highlighted in the space-time cube visualization. Figure 3.7 illustrates this feature.

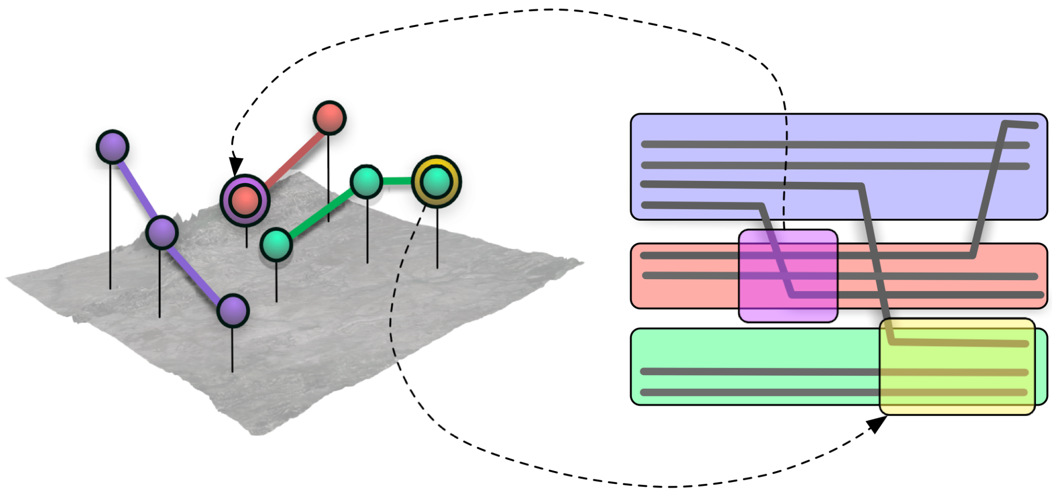

Figure 3.7: Combining Affiliation Timeline and space-time cube within one visualization environment. Semantic cross-highlighting causes data to be selected in one depiction to be automatically highlighted in the other.

3.4 Summary

This chapter proposed a methodology for the visual analysis of dynamic social networks that is useful for both sociologists interested in studying the structures of various social systems, as well as other domain scientists who also seek to understand the environmental factors that affect the social interaction. SocioScape combines a novel visual representation method, the Affiliation Timeline, with the well established space-time cube technique for visualization of spatio-temporal data. The Affiliation Timeline depicts the evolution of social groups, as well as the association choices actors make over time. By eliminating edges and nodes present in graphs, the Affiliation Timeline could provide an intuitive mechanism for an analyst to explore the social structure at all timesteps. Furthermore, preserving the position of actors and communities within the diagram across all timesteps makes the diagram more readable than traditional graph representations. The next chapter presents an interactive visualization environment that implements the proposed methodology.

4. IMPLEMENTATION

This chapter describes the design and implementation of SocioScape, an interactive visualization environment that allows an analyst to explore one or more dynamic social network datasets. The visualization environment implements the pipeline presented earlier and supports the Affiliation Timeline as well as the space-time cube visualization.

4.1 Programming environment

The visualization environment was implemented in Electro (62). Electro is a run-time application development environment designed for real-time 3D and 2D graphics with hardware acceleration. There are a number of advantages that make Electro an attractive option:

- Rapid implementation: programming in Electro is done using the Lua, a simple and highly flexible scripting language. This allows rapid development of interactive visualization applications that can be easily extended in a short amount of time.

- Platform independence: Electro runs on all major operating systems, including Windows, MacOS X, and Linux. It also supports a wide variety of architectures including 32-bit and 64-bit Intel and PowerPC.

- Support for tiled display: Electro supports a wide variety of display platforms ranging from single screen laptops to scalable tiled displays that have resolutions on the order of hundreds of millions of pixels. This means that SocioScape can be used on desktop computers with single displays, as well as wall sized tiled displays (60).

4.2 User interface

The user interface has been designed primarily to allow users to take advantage of the available screen space by arranging different depictions of the data side-by-side for cross-analysis and correlation. Additionally, the program allows multiple different datasets to be visualized simultaneously for comparison. Figure 4.1 illustrates the user interface.

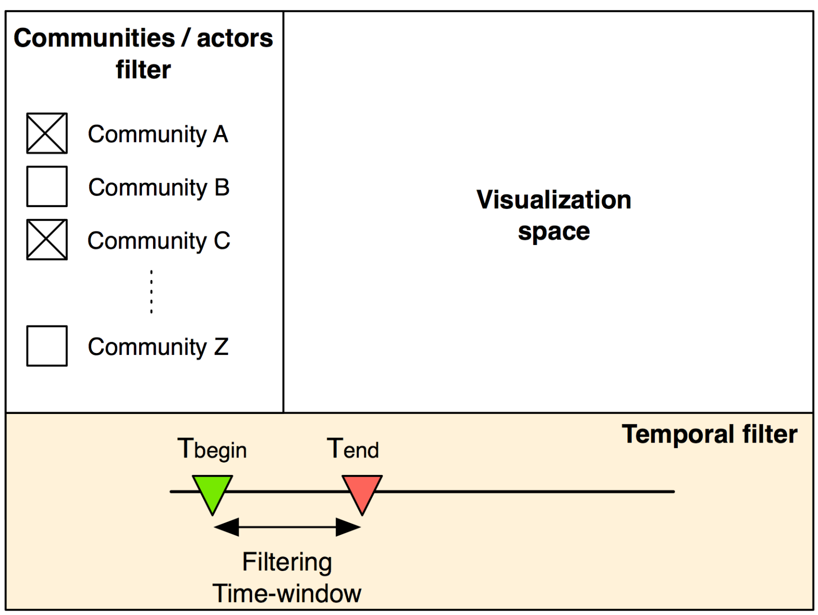

Figure 4.1: User interface. The bottom area contains two sliders for specifying the filtering time-window. Interactions that falls outside the specified temporal filter do not show up in the visualization. The sidebar on the left allows selecting a subset of actors or communities to be visualized. The while area is where the actual visualization gets rendered.

The visualization space displays one of the two depiction modes; the space-time cube or the Affiliation Timeline. The depiction mode can be switched from a pop-up menu that can be brought up by right-clicking anywhere in the visualization space. The menu also allows switching between the different available datasets.

The interface allows an analyst to interactively modify the filters, causing the visualization to be re-rendered. The temporal filter allows the user to specify a continuous time-window that falls anywhere within the observation period. The length and position of the time-window can be adjusted using two sliders: the green slider specifies the beginning timestep (T_{begin}), whereas the red slider specifies the ending timestep (T_{end}). Interactions and community sightings the occur outside the filtering time-window are not shown in the visualization. On the left side of the screen, the communities/actors filter panel allows the user to select a subset of communities or actors to be depicted in the visualization. This panel can be retracted to enlarge the area available for the visualization.

To allow the analyst to visually compare and correlate the different depictions and datasets, the visualization space can be divided into a number of smaller spaces. Figure 4.2 illustrates this concept. Each visualization space can independently show one of the two depiction modes with its own set of filters. The user has complete control over the arrangement and sizes of these spaces. This added control affords more flexibility to spatially arrange the relevant views to ease the task of comparing and correlating the different depictions. Figure 4.3 is screenshot of the application showing both the space-time cube and the Affiliation Timeline. Figure 4.4 shows another arrangement in which the recorded group movement of two populations of wild animals is shown side-by-side.

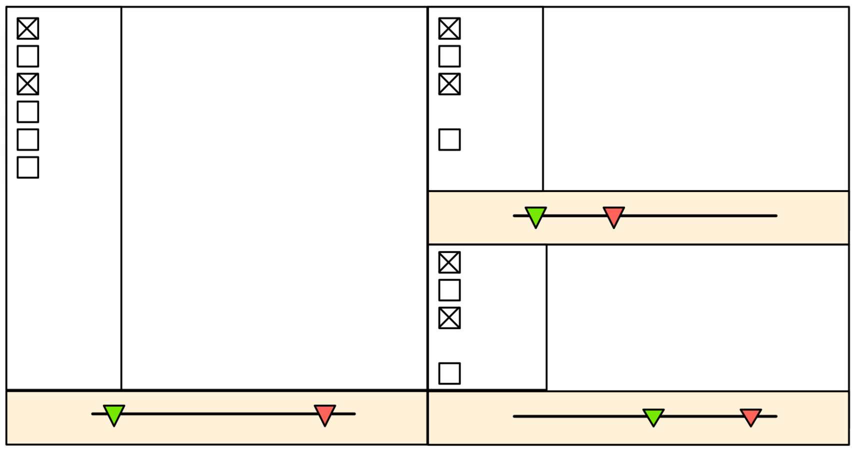

Figure 4.2: The visualization space can be divided into smaller spaces with each space showing one of the two depiction modes with its own set of filters.

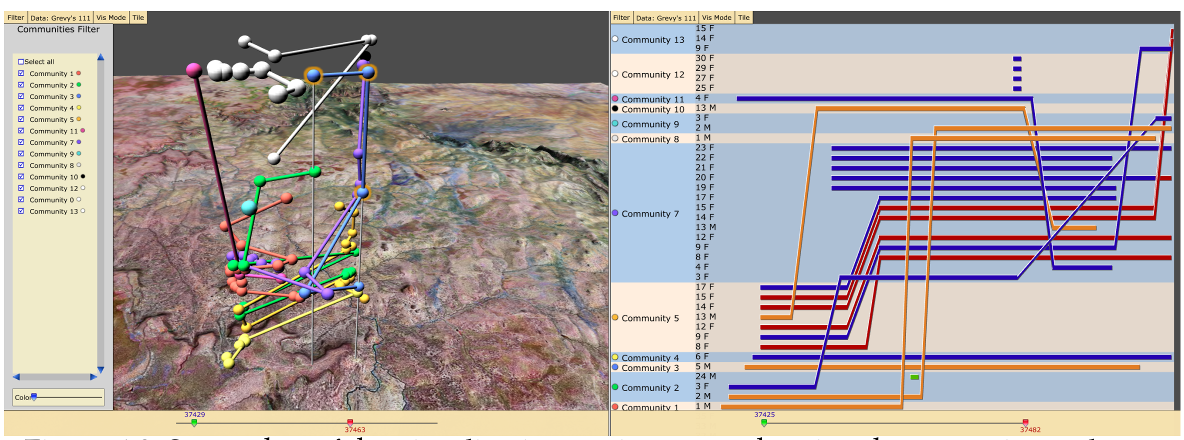

Figure 4.3: Screenshot of the visualization environment showing the space-time cube on the left and the Affiliation Timeline on the right.

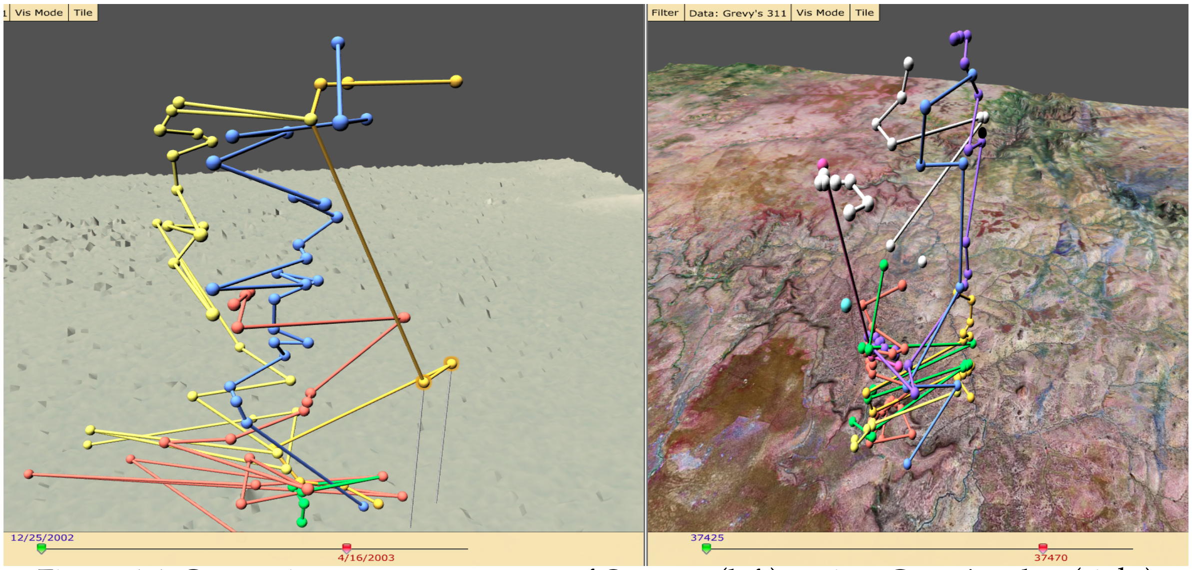

Figure 4.4: Comparing group movement of Onagers (left) against Grevy's zebra (right)

4.3 Interaction

To interact with the visualization space, there are a number of controls that apply to both the Affiliation Timeline and space-time cube:

- Zooming: the view can be zoomed in/out using the mouse wheel. An analyst can zoom out to see an overview of the interactions and attempt to look for an overall pattern. Conversely, zooming in allows one to focus on a smaller portion of the dataset.

- Translation: When the view is zoomed in, not all the data might fit in the available visualization space. To show the hidden data, the user can simply translate the view by dragging the mouse while holding the right button. Since both the Affiliation Timeline and space-time cube unrolls the temporal dimension linearly in space on the Z and X axes, respectively, translating the visualization in one of the two directions simulate the passage of time. On the other hand, translating the view in the other dimension allows one to explore the different actors and communities.

- Highlighting: The affiliation history of an actor in the Affiliation Timeline can be highlighted by clicking the line that depicts the actor of interest. This causes the other actors in the visualization to be de-emphasized by decreasing the opacity of their lines. Similarly, a community sighting can be highlighted by clicking on its sphere in the space-time cube, causing it to glow.

4.4 Semantic Cross highlighting

Although the different visualization spaces shown simultaneously can be independently manipulated to show different portions of the dataset at different time periods, the environment supports a shared highlighting context. When the user highlights data in one visualization space, the relevant data in all other spaces is automatically highlighted. For example, when a community sighting is highlighted in the space-time cube, a portion of that community's rectangle is automatically highlighted in the Affiliation Timeline. This highlighted portion shows actors affiliated with that community at the time of sighting. Figure 4.5 shows an example of this.

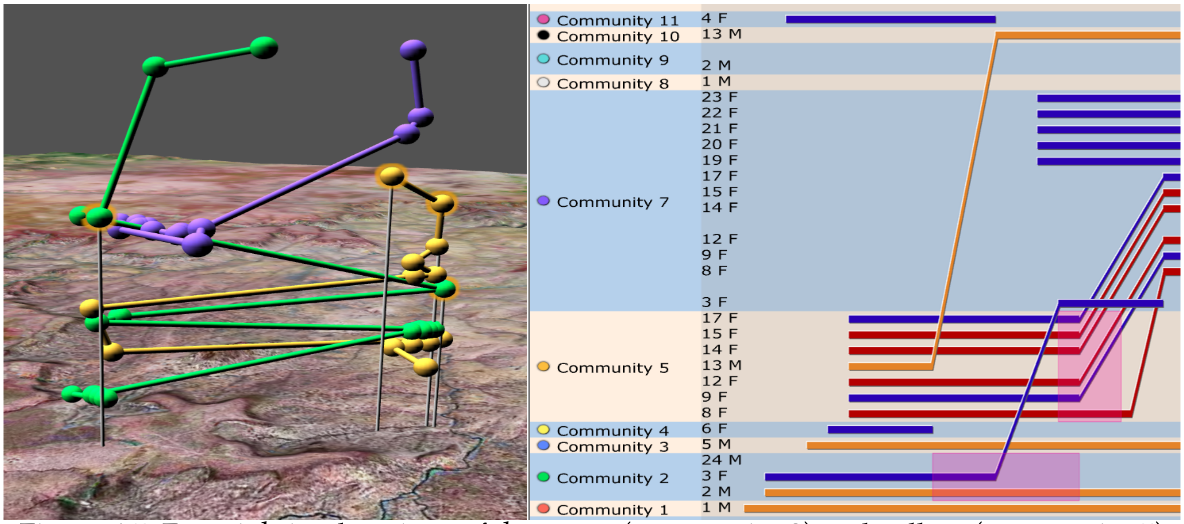

Figure 4.5: Few sighting locations of the green (community 2) and yellow (community 5) selected in the space-time cube (left) with the corresponding social structure highlighted in the Affiliation Timeline in pink (right)

4.5 Display scalability

One of the novel features in the Affiliation Timeline is that it projects the complex social interactions that happen over time onto a planar diagram. This allows the visualization to scale well as the number of timesteps increases; the width of the diagram can be simply increased to allow interactions over longer time periods to be depicted. However, the visualization will always be limited by the available screen resolution.

The use of scalable tiled-displays is becoming increasingly common in research environments (60). These displays overcome the limited resolution of a single displays by simply tiling up multiple displays, creating wall-sized displays that can be interactively used to display more informations. Our implementation of SocioScape takes advantage of such platforms to increase the available screen resolution available to the application. This allows one to visualize longer periods of interactions. Alternatively, a larger number of datasets can be displayed side-by-side.

5. CASE STUDY

This chapter presents an evaluation of SocioScape for the exploration of real-world, spatially referenced, dynamic social network datasets. We present a case study in which the implementation described earlier is used by a team of ecologists to explore two datasets depicting the grouping behavior and movement of two populations of endangered species. The first population belongs to the Grevy's zebra specie, found in northern central Kenya. The second population is a member of Onagers, a specie of wild asses that is similar in appearance to donkeys. This population reside in the eastern deserts of India. The two species are in danger of extinction due to a variety of environmental factors as well as human-related activities in their habitat.

There is a growing interest among behavioral ecologists in the social behavior of these species. Ecologists seek to understand how these populations are responding to environmental challenges such as predation risk and dwindling food resources, and how their social interaction is being affected by these factors. This information allows ecologists to devise appropriate conservation plans to provide protection to the remaining populations.

SocioScape can be effectively used to explore potential answers to some of these questions. Using the space-time cube and the Affiliation Timeline side-by-side, an ecologist can see how herds or groups of animals move in the environment, and simultaneously look at the composition of these social groups at the time they were sighted. Conversely, an ecologist can look at the Affiliation Timeline for interesting phenomena (for example, an individual switching affiliation from one community to another), highlight the appropriate portion in the diagram, and look at the geographical location where this phenomenon occurred. This could give clues on the influence of various environmental factors on the behavior of these animals.

To illustrate the advantages of using SocioScape in this domain, we first describe the visualization methods that the ecologists currently use. The limitations of these techniques are pointed out, along with a discussion of how SocioScape addresses them. Follows this is a presentation of a user study that was conducted to evaluate SocioScape as an exploratory tool with the participation of experts in the field of behavioral biology. Throughout the user study, the visualization environment described earlier is used to explore two datasets comprising observations of social interaction and movement of Grevy's zebras and Onagers.

5.1 Datasets

The datasets used in this case study were obtained by observing animal groups in the wild and recording their location, as well as the individuals that were present in these groups. The identification of individuals in Grevy's zebra groups was done by analyzing the striping patterns on the animal's skin (each zebra individual has a unique pattern). For each group sighting, three pieces of information was recorded:

- The GPS coordinates of the sighting location.

- The date and time of the sighting, which we refer to as the timestep.

- Ids of the individual that were present in the group. All the individuals present in the group are considered to be interacting with each others.

The Grevy's zebra dataset contains social interaction and movement of a Grevy's zebra population in the Mpala conservation ranch in northern central Kenya. The numbers of individuals in the dataset is 35. The animals were observed in a 45 square kilometers area over a period of two months. During this period, 149 sightings of zebra groups were made. The Onager's dataset comprises 41 individuals. The number of group sightings is approximately 350 spread over a 75 square kilometers area over a period of approximately 6 months.

5.2 Earlier techniques

Earlier visualization techniques used by behavioral ecologists depicted either the movement of individuals, or the social structure of the population. These depictions were generated and rendered separately, and could not be easily integrated into one interactive environment.

5.2.1 Visualization of community structure

To visualize the dynamic community structure of the population, the ecologist used a graph-variant proposed in (3). Figure 5.1 shows the community affiliation in Grevy's zebra visualized using this technique. Figure 5.2 shows a visualization of the same dataset using the Affiliation Timeline.

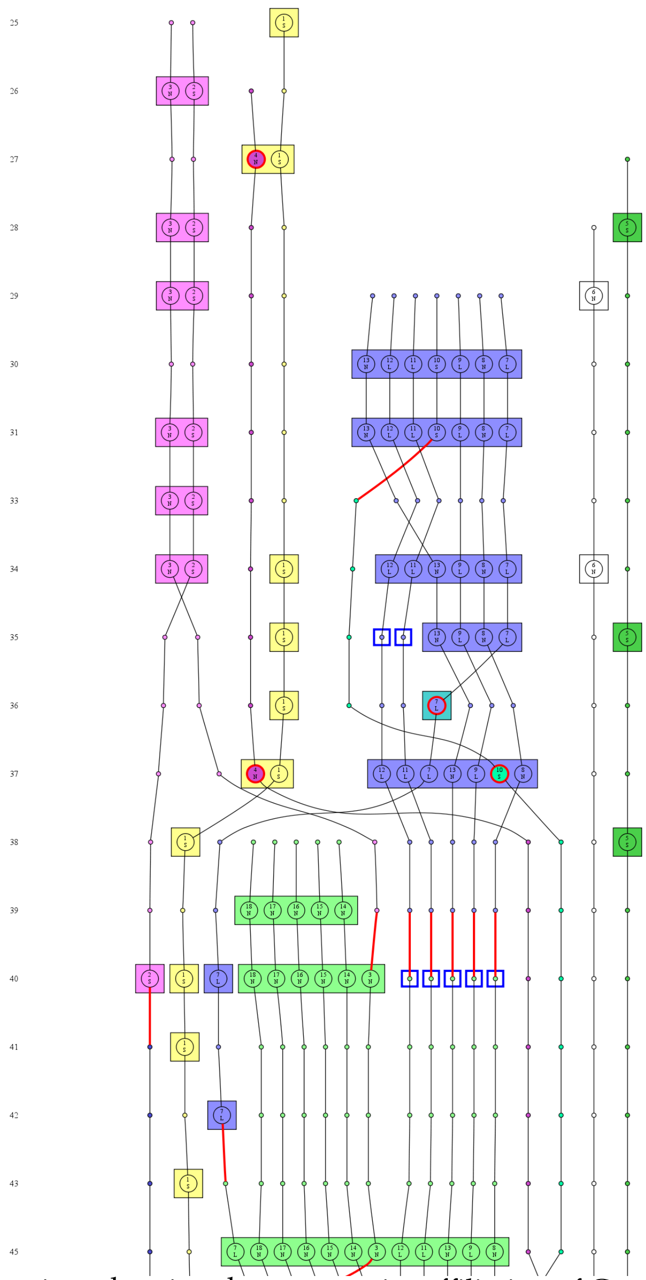

Figure 5.1: A graph-variant showing the community affiliation of Grevy's zebra in the first 19 timesteps. Individuals are depicted with nodes and communities are depicted with colored rectangles encompassing individuals affiliated with those communities. Every node has exactly one incoming and one outgoing edge that link it to its two siblings representing the same individual in the previous and next timesteps, respectively. Circles indicate that an individual was present when the sighting of the community was made. A dot indicates that an individual was missing. Red edges depict an individual switching affiliation from one community to another. Source: (3)

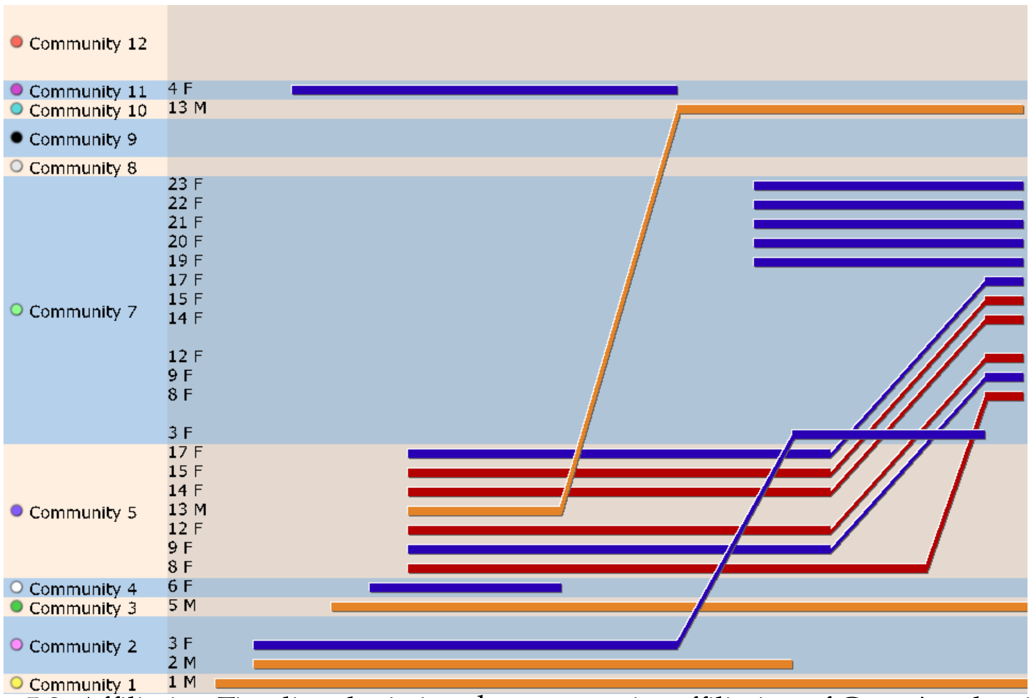

Figure 5.2: Affiliation Timeline depicting the community affiliation of Grevy's zebra in the first 19 timesteps. Lines depict individuals as they change affiliation from one community to another. The color of the line indicates the reproductive state of the animal; Orange for stallion, Red for lactating female, Blue for non-lactating female, and Green for bachelor.

One of the problems in diagram 5.1 is that although edges help trace an animal's community affiliation history, this still requires significant effort from the analyst, making the diagram difficult to interpret. The Affiliation Timeline solves this problem by getting rid of nodes and edges, and depicting an animal's affiliation history with one solid line. For example, in figure 5.1, it is difficult to notice that individual number 3 switches its affiliation from the pink to the green community. In contrast, this switch is very easy to notice in figure 5.2. Another problem with the layout in figure 5.1 is that it is not stable. That is, the position of individuals and communities in the layout changes for no obvious reason. In contrast, the position of individuals in the Affiliation Timeline is stable as long as their affiliation does not change. The positions of the different communities is also stable across all timesteps.

The diagram shown in figure 5.1 also contains the reproductive state for each animal at all timesteps. This is depicted with a single character inside the individual's node (S for stallion, L for lactating female, N for non-lactating female, and B for bachelor). Unfortunately, this does not help an analyst monitor the reproductive status of an individual over time because this requires sequentially scanning the individual's nodes at each timestep. By depicting this information with color (Orange for stallion, Red for lactating female, Blue for non-lactating female, and Green for bachelor), the Affiliation Timeline makes it easier to notice a change in the reproductive state of an animal as it becomes evident when the solid line changes color.

A limitation of the Affiliation Timeline is that it does not make an indication when an individual is missing from its community. Often, when observations of animal groupings are recorded, some individuals might not be present with their community. This is features could be incorporated in future implementations. One possible way to depict an individual's absence from its community is to dim the color of its line at the time the individual was missing.

5.2.2 Visualization of individual movement

Understanding movement patterns of animals is essential to understanding their social behavior. In fact, a great deal of the social behavior exhibited by animals (including Grevy's zebras and Onagers) is directly influenced by the need of individuals or groups to move in search for food and water, among many other factors.

With the advent of sensor networks, it is becoming easier to track the movement of individuals and groups. This data has been traditionally visualized on a 2D map of the habitat with linear tracks connecting two consecutive sighting locations. The ecologists have been using an interactive variant of this technique that animates the movement of individuals over the landscape (61). Figure 5.3 shows a screenshot of this visualization.

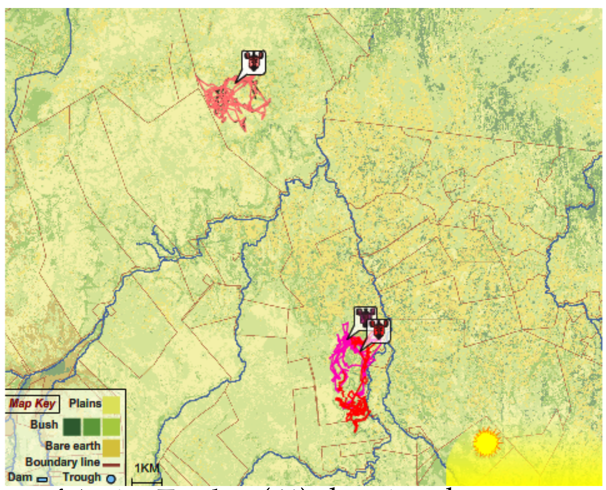

Figure 5.3: Screenshot of ActionTracker (61) showing the movement trajectory of three zebras. Tracks are coded with unique colors to differentiate the individuals.

A problem with this approach is that it is susceptible to cluttering. This makes it harder for an observer to trace the movement of animals. Furthermore, this technique uses animation to show the time period it took an individual or a group to move from one point which only allows an analyst to perceive that period for the last few frames of the animation.

A major disadvantage of using two separate environments to visualize the social structure and the physical movement of animals is that there is no way of relating observation in one visualization to the other. For instance, if an observer finds that a particular individual suddenly made an unexpected movement at a particular moment in time, there observer will not be easily able to investigate the social structure of the population at that moment.

The tight integration of the spatio-temporal visualization, which shows the movement of groups with the Affiliation Timeline, which shows the community structure of the population allows an ecologists to investigate how the movement of communities and individuals give rise to the underlying social structure. Consequently, the different ecological circumstances that influence the social behavior of these animals could be better understood. This information in turn could be used to determine appropriate habitat that could be later set aside for the conservation.

5.3 User study

In order to validate the effectiveness of SocioScape for the exploration of spatially referenced dynamic social networks, a user study in has been conducted in which behavioral ecologists used the implementation described earlier to explore the social behavior and movement of Grevy's zebras and Onagers. The participants in the user study were expert researchers who originally collected both datasets. Hence, they were familiar with the overall social behavior of the two species, as well as their habitats. Furthermore, the participants were motivated to test the new system to further develop their understanding of the two datasets for their own research agenda.

The user study followed traditional software evaluation trials. The trial had a number of participants simultaneously use the visualization environment to explore the Grevy's zebra and Onager dataset. The visualization was rendered on a tiled-display consisting of two 30-inch LCD displays placed side-by-side. This allowed multiple users to simultaneously look at the same data and discuss their conclusions and inferences among themselves. Figure 5.4 shows the setup.

Figure 5.4: Users utilizing a dual 30-inch tiled-display setup of SocioScape to analyze the movement and grouping behavior of Grevy's zebras

The number participants in the user study was 10, 4 of them were ecologists, and the rest were computer scientists. Before beginning the trial, the participants received a brief training covering the use of the visualization program and its features. The participants were not asked to perform any specific task. Rather, they were encouraged to freely interact with the visualization to explore the datasets and discuss the inferences among themselves. The trial lasted for approximately 2 hours. During the trial, the participants were video and audio recorded. These recordings were later analyzed to determine if the visualization was effective and whether it helped expand the participants knowledge of the datasets.

5.4 Findings

The participants started by dividing the visualization space into two smaller spaces; the first showed the space-time cube visualization on one display, while the other showed the Affiliation Timeline on the second display.

At first, the participants started looking at the Affiliation Timeline to get a sense of the overall community structure of the population, and to look for interesting communities (such as unusually large ones). At this point, the participants would apply a filter to the space-time visualization to show a subset of the detected communities which they deemed interesting. Then the participants started to look at how these communities moved in the landscape. While tracing the movement of communities, the users would highlight one or more of the sighting points. This would cause the social composition of these communities to be automatically highlighted in the Affiliation Timeline, allowing the ecologists to inspect the individuals affiliated with those communities at the time of sighting. Conversely, the participants would also look for interesting observations in the Affiliation Timeline (such as a group of individuals changing their community affiliation at once) and investigate the physical position at which these events occurred.

The ecologists appreciated the fact that they can simultaneously see the movement of communities as well the social structure of the population at different points in time. In particular, Dr. Daniel Rubenstein, chair of the department of Ecology and Evolutionary Biology at Princeton University, responded with the following comments.

> “We are learning about what motivates individuals and how that affects the dynamics of the group”. “This is very useful at showing us how individuals and communities are behaving”

Occasionally, the participants would notice an unusual movement in the space-time visualization, such as a community moving a significant distance in a short amount of time. This would trigger them to highlight the sighting location at which this movement was recorded, and look at the Affiliation Timeline to see if there were any changes to the community structure at that point.

The combination of space-time cube and Affiliation Timeline has revealed some discoveries that would have been otherwise very difficult to see. For instance, two of the communities attracted a participant's attention after looking at their movement in the space-time cube. These two communities seemed to be oscillating periodically between two sites in the landscape. The ecologists identified these two sites as a grazing and a drinking site. Although these communities were sharing the two sites and moving between them periodically, there movement was one day apart. Furthermore, looking at the Affiliation Timeline confirms that these communities remain separate for some time before they each joined a third community. Figure 5.5 illustrates this phenomenon. This led to the hypothesis that these two communities were avoiding each others before the emergence of the third community:

> “Even though these communities lineup in space very close to each others, they are off by a day. This is very robust at telling us they are avoiding each others.”

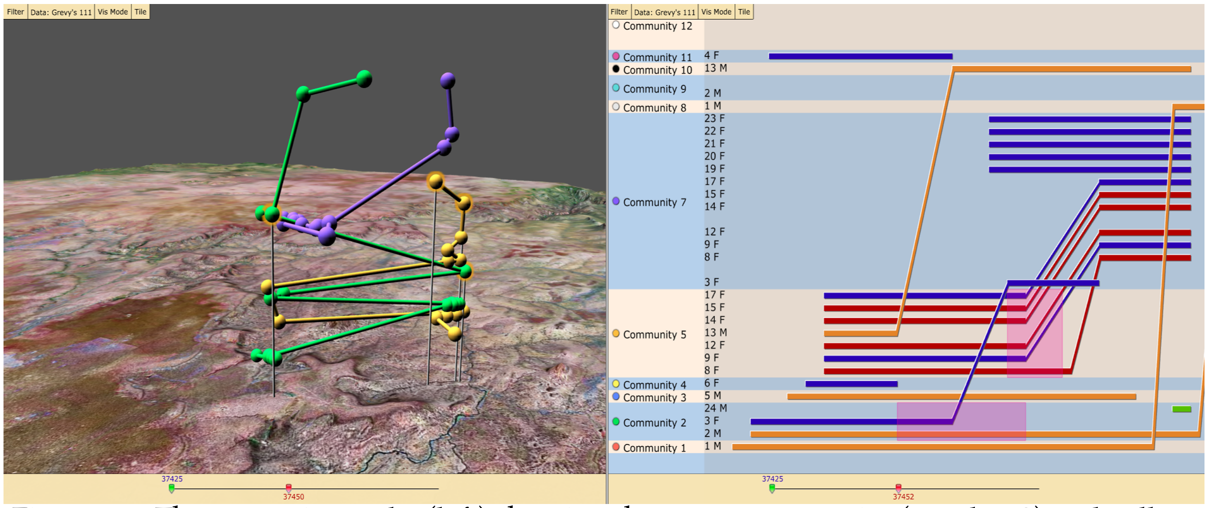

Figure 5.5: The space-time cube (left) showing the green community (number 2) and yellow community (number 5) oscillating periodically between two sites with. However, their periodic movement is one day apart in time. The highlighted portion of the Affiliation Timeline (right) shows later that members of these two communities eventually merge into the purple community (number 7).

The participants also liked the fact that the Affiliation Timeline shows the community affiliation of individuals along with their reproductive state:

> “We are looking at a different projection that shows the individual by state moving in and out of the community, which is very useful.”

The participants also found it easy to trace the community affiliation of individuals over time by looking at the Affiliation Timeline. This provided some evidence that the representation method used in the Affiliation Timeline is more intuitive than graphs:

> “It is easier to see the individuals move. That was really very easy.”.

> “This is a very clean depiction of community membership.”

The usage pattern as well as the findings that the participants were able to infer after using SocioScape validate the hypothesis that the combination of the space-time cube with the Affiliation Timeline is useful for investigating the spatial patterns, and how the physical decision making of individuals give rise to the underlying social structure. It also supports the claim that the Affiliation Timeline allows easy perception of the community structure, and how this structure changes over time.

Overall, The reaction to SocioScape was very positive. The participants acknowledge the usefulness of integrating a spatio-temporal visualization with a depiction that illustrates the social interaction in a single visualization environment. This combination allowed the ecologists to study how the micro-dynamics of the population embodied by the movement of communities and individuals give rise to the underlying social structure. Dr. Rubenstein provided the following comments:

> "[This visualization] finally put time and space together. This allows us to understand the physical decision making that lead to the shaping of communities. The dynamic community analysis gave us a better picture for understanding zebra dynamics. The space will give us even a better picture of that temporality."

> "This is giving us the spatial structure, and that's the power of this visualization."

> "What this [visualization] allows us is infer some of the dynamics when we were not there."

6. CONCLUSION AND FUTURE WORK

This thesis presented SocioScape, an interactive visualization tool that embodies a methodology for the visual analysis of spatial and temporal group dynamics in social networks. The methodology introduces a novel visual representation technique suitable for dynamic social networks. This representation provide an advantage over dynamic graphs by explicitly illustrating the evolution of social groups and association choices made by actors over time. The representation is combined with a well-established technique for depicting spatio-temporal data, allowing analysts to investigate the effect of the physical positioning of actors and their movement in the environment on their social behavior. This integration also facilitates the investigation of potential hypotheses that explain the emergence of the observed social structure structure.

6.1 Contributions

Earlier work on the visualization of dynamic social networks was inspired by automatic graph layout algorithms. While these techniques work well for static social networks, they suffer from a number of limitations when applied to dynamic networks that change with time. Furthermore, while these techniques provided overview of the social structure, they are largely useless at explaining why this structure developed in the first place. While sociologists are usually interested in the abstract social interactions between actors, other domain scientists seek to understand the role of the environment in shaping that interaction.

SocioScape addresses these limitations by proposing a novel visual representation for the depiction of social interaction. These representations radically depart from traditional graph inspired representations, focusing on revealing group membership choices that actors make over time. The depiction exploits properties of the human visual processing system to create an easily perceptible representation of the social structure. Furthermore, SocioScape provides a spatio-temporal visualization to represent the movement of groups and actors. This provides opportunities for investigating the role of external environmental factors and the positioning of actors within the environment on shaping the dynamics of interactions.

The applicability of SocioScape is demonstrated with a case study in which expert behavioral ecologists utilized an interactive implementation of the methodology to explore the social behavior and movement of two populations of endangered species. The results of the case study provided tangible example and positive support that confirm the effectiveness of the methodology.

In summary, the contributions of this thesis are:

1. A novel visual representation method for dynamic social networks. The technique departs from traditional graph-based visualizations, employing depictions that are easier to interpret, revealing the evolution of social groups and association choices that actors make over time.

2. A methodology that integrates abstract representations of social interactions with a spatio-temporal visualization. This integration supports the investigation of the role

of environment in shaping the underlying social structure, and allows opportunities for exploring potential answers as to why that structure developed in the first place.

3. A case study in which the methodology was used to explore the social behavior and the physical movement of two populations of endangered species, validating the usefulness of the methodology.

6.2 Future work

The tight integration of spatial and temporal analysis in SocioScape proved to be very useful in exploring potential explanations for the emergence of certain social structures as a function of group/actor movement and external environmental influences. While the spatial analysis model proposed by SocioScape is currently limited to geographically-bounded social interactions, this model could be expanded to analyze other kinds of interactions, and explore other types of external influencing factors that are not necessarily grounded in the physical positioning of actors within the environment. For this to be happen, a mapping model that projects different types of environmental attributes onto a (spatial) visualization needs to be developed.

The user study suggests that the Affiliation Timeline is easier to interpret and understand than traditional graph layouts. However, more experimentation needs to be done so that this claim can be statistically proven. Additionally, more testing needs to be done in order to verify the scalability of the diagram in terms of the number of actors/groups, and the length of the observation period.

The Affiliation Timeline could also use a number of improvements. First, a layout algorithm that minimizes line crossings could greatly enhances the readability of the diagram. Second, the user interface can be improved to include functionality to reduce the amount of information without having to resort to filters. For example, the rectangles representing communities could be collapsed or expanded on demand to show the composition of communities, without significantly disrupting the layout. Lastly, the diagram should be expanded to depict additional notions that the dynamic community identification framework supports, such as the notion of a visiting actor (an actor observed interacting with a different community, nonetheless retaining affiliation with its original community).

CITED LITERATURE

1. Freeman, L. C.: The Development of Social Network Analysis: A Study in the Sociology of Science. Empirical Press, 2004

2. Wasserman, S. and Faust, K.: Social Network Analysis: Methods and Applications (Structural Analysis in Social Science). Cambridge University Press, 1994

3. Tantipathananandh, C., Berger-Wolf, T., Kempe, D.: A framework for community identification in dynamic social networks. In KDD '07: Proceedings of the 13th International Conference on Knowledge Discovery and Data Mining, San Jose, CA, USA, 2007. ACM Press

4. Carley, K. M.: Dynamic Network Analysis. In the Summary of the National Research Council workshop on Social Network Modeling and Analysis, Breiger, R. and Carley, K. M. (Eds.). Washington, DC, 2003. The National Academy Press

5. Fischhoff, I. R., Dushoff, J., Sundaresan. S. R, Cordingley, J. E., Rubenstein, D. I.: Reproductive Status Influences Group Size and Persistence of Bonds in Male Plains Zebra (Equus burchelli). Behavioral Ecology and Sociobiology, 63(7):1035-1043, 2009

6. Sethi, A., Eargle, J. Black, A. A., Luthey-Schulten, Z.: Dynamical networks in tRNA: protein complexes. In Proceedings of the National Academy of Sciences of the USA, 106(16):6620-6625, 2009

7. Krause, J., Lusseau D., James, R.: Animal social networks: an Introduction. Behavioral Ecology and Sociobiology, 63(7):967-973, 2009

8. Anderson, R. M.: Epidemic models and social networks. Math. Scientist, 24:128-147, 1999

9. Hethcote, H. W.: The mathematics of infectious diseases. SIAM Review, 42:599-643, 2000

10. Girvan, M. and Newman, M.: Community structure in social and biological networks. In Proceedings of the National Academy of Sciences, 99(12):7821-7826, 2002

11. Newman, M.: The structure and function of complex networks. SIAM Review, 45:167-256, 2003

12. Moreno, J. L.: Who Shall Survive?. Beacon House, Beacon, NY, 1953. Available from: http://www.asgpp.org/docs/WSS/WSS.html

13. Battista, G. D., Tamassia, R., Tollis, I.: Algorithms for drawing graphs: an annotated bibliography. Computational Geometry-Theory and Applications, 4(5):235-282, 1994

14. Fruchterman, T. M. J. and Reingold, E. M.: Graph Drawing by Force-Directed Placement. Software: Practice and Experience, 21(11), 1991

15. Freeman, L. C.: Visualizing social networks. Journal of Social Structure, 1, 2000. Available from: http://www.cmu.edu/joss/content/articles/volume1/Freeman.html

16. Bender-deMoll, S. and McFarland, D. A.: The Art and science of dynamic network visualization. Journal of Social Structure, 7(2), 2006. Available from: http://www.cmu.edu/joss/content/articles/volume7/deMollMcFarland/

17. Batagelj, V. and Mrvar, A.: Pajek: Program for Large Network Analysis. http://vlado.fmf.uni-lj.si/pub/networks/pajek/

18. NWB Team: Network Workbench Tool. Indiana University, Northeastern University, and University of Michigan, 2006. http://nwb.slis.indiana.edu

19. Heer, J. and Boyd, D.: Vizster: Visualizing online social networks. In InfoVis '05: Proceedings of the 2005 IEEE Symposium on Information Visualization, Minneapolis, MN, USA, 2005. IEEE Press

20. Adar, E.: GUESS: a language and interface for graph exploration. In CHI '06: Proceedings of the 2006 SIGCHI conference on Human Factors in computing systems, Montréal, Québec, Canada, 2006. ACM Press

21. Suh, B., Chi, E. H.: Us vs. them: understanding social dynamics in Wikipedia with revert graph visualizations. In VAST '07: IEEE Symposium on Visual Analytics Science and Technology, Sacramento, CA, USA, 2007. IEEE Press

22. Graphvis – Graph visualization Software. Available from: http://www.graphviz.org/

23. Baumes, J., Goldberg, M., Magdon-ismail, M., Wallace, A.: Discovering Hidden Groups in Communication Networks. In in Proceedings of the 2nd NSF/NIJ Symposium on Intelligence and Security Informatics, Tucson, AZ, USA, 2004. Lecture Notes in Computer Science 3073, Springer, 2004

24. Battista, G. D., Eades, P., Tamassia, R., Tollis, I., G.: Graph drawing: algorithms for the visualization of graphs. Prentice Hall, 1999

25. Auber, D., Chricota, Y., Jourdan, F., Melancon, G.: Multiscale visualization of small world networks. In InfoVis '03: IEEE Symposium on Information Visualization, Seattle, WA, USA, 2003. IEEE Press

26. Milgram, S.: The small world problem. Psychology Today, 2:60-67, 1967

27. Ham, F. and Wijk, J.: Interactive Visualization of Small World Graphs. In InfoVis '04: 2004 IEEE Symposium on Information Visualization, Austin, TX, USA, 2004. IEEE Press

28. Fisher, D. and Dourish, P.: Social and Temporal Structures in Everyday Collaboration. In CHI '04: Proceedings of the 2006 SIGCHI conference on Human Factors in computing systems , Vienna, Austria, 2004. ACM Press

29. Jeong, H., Mason, S. P., Barabasi, A., Oltavai, Z. N.: Lethality and centrality in protein networks. Nature, 411:41-42, June 21, 2001

30. McPherson, J., Ma, K., Ogawa, M.: Discovering parametric clusters in social small-world graphs. In SIGAPP '05: 2005 ACM Symposium on Applied Computing, Socorro, NM, USA, 2005. ACM Press

31. Heer, J.: Exploring Enron: Visual data mining of E-mail. Available from: http://jheer.org/enron/

32. Kershenbaum, A. and Murray, K.: Visualization of network structures. Journal of Computing Sciences in Colleges, 21(2):59-71, 2005

33. Jia, Y., Hoberock, J., Garland, M., Hart, J. C.: On the visualization of social and other scale-free networks. IEEE Transactions on Visualization and Computer Graphics, 14(6):1285-1292, 2008

34. Branke, J.: Dynamic graph drawing. Drawing Graphs, volume 2025 of Lecture Notes in Computer Science, pages 228-246. Springer-Verlag, 2001

35. Kumar, G. and Garland, M.: Visual Explication of Complex Time-varying Graphs. IEEE Transactions on Visualization and Computer Graphs, 12(5):805-812, 2006

36. Yang, X. Asur, S., Parthasarathy, S.: A visual-analytics toolkit for dynamic interaction graphs. In KDD '08, Las Vegas, Nevada, USA, 2008. ACM Press

37. Falkowski, T., Bartelheimer, J., Spiliopoulou, M.: Mining and visualizing the evolution of subgroups in social networks. In IEEE/WIC/ACM International Conference on Web Intelligence, Hong Kong, 2006. IEEE Press

38. Brandes, U., Corman, S., R.: Visual unrolling of network evolution and the analysis of dynamic discourse. Information Visualization, 2:40-50, 2003

39. Gaertler, M. and Wagner, D.: A Hybrid model for drawing dynamic and evolving graphs. Graph Drawing, volume 3843 of Lecture Notes in Computer Science, pages 189-200. Springer-Verlag, 2006

40. Kang, H., Getoor, L., Singh, L.: Visual Analysis of Dynamic Group Membership in Temporal Social Networks. ACM SIGKDD Explorations Newsletter, 9(2):13-21, 2007

41. Viégas, F. B., Boyd, D., Nguyen, D. H.: Digital artifacts for remembering and storytelling: PostHistory and Social Network Fragments. In Proceedings of the 37th Hawaii International Conference on System Sciences, Waikoloa, Hawaii, USA, 2004. IEEE Press

42. Viégas, F., Wattenberg, M., Dave, K.: Studying cooperation and conflict between authors with history flow visualizations. In CHI '04, Vienna, Austria, 2004

43. Viégas, F., B. and Donath, J.: Social network visualization: can we go beyond the graph?. In CSCW '04 : Workshop on Social Networks, Chicago, IL, USA, 2004. ACM Press

44. Blythe, J., McGrath, C., Krackhardt, D.: The effect of graph layout on inference from social network data. Proceedings of the Symposium on Graph Drawing, volume 1024 of Lecture Notes In Computer Science, pages 40-51. Springer-Verlag, 1995

45. Purchase, H.: Which aesthetic has the greatest effect on human understanding?. Graph Drawing, volume 1353 of Lecture Notes in Computer Science, pages 248-261. Springer-Verlag, 1997

46. Purchase, H., C.: Effective information visualization: a study of graph drawing aesthetics and algorithms. Interfacing with Computers, 13(2):147-162. Elsevier, 2000.

47. Purchase, H. C., Hoggan, E., Görg, C.: How important is the "Mental Map"? - An Empirical Investigation of Dynamic Graph Layout Algorithm. In GD '06: 14th International Symposium on Graph Drawing, Karlsruhe, Germany, 2006. In Lecture Notes in Computer Science 4372, Springer, 2007

48. Saffrey, P., Purchase, H.: The "mental map" versus "static aesthetic" compromise in dynamic graphs: a user study. ACM International Conference Proceedings Series, 314:85-93. ACM, 2008

49. Friedrich, C., Eades, P.: Graph drawing in motion. Journal of Graph Algorithms and Applications, 6(3):353-370, 2002

50. Fayyad, U. M., Grinstein, G. G.: Introduction. In Information Visualization in Data Mining and Knowledge Discovery, Morgan Kaufmann, Los Altos, CA, 2001, pp. 1–17

51. Hägersrand, T.: What about people in regional science? Papers of the Regional Science Association, 24:6-21, Springer-Verlag, 1970

52. Andrienko, N., Andrienko, G., Gatalsky, P.: Impact of data and task characteristics on design of spatio-temporal data visualization tools. In Exploring Geovisualization. (Eds: Dykes, J.A., Kraak, M.J., and MacEachren, A.M.), pages: 202-222. Elsevier, London, 2005

53. Karrak, M.: The space-time cube revisited from a geovisualization perspective. In Proceedings of the 21st International Cartographic Conference, 2003

54. Andrienko, G., Andrienko, N., Wrobel, S.: Visual analytics tools for analysis of movement data. SIGKDD Explorations Newsletter, 9(2):38-46, 2007

55. Andrienko, N., Andrienko, G., Gatalsky, P.: Exploratory spatio-temporal visualization: an analytical review. Journal of Visual Languages and Computing, 14(6):503-541. Elsevier, 2003

56. Eccles, R., Kapler, T., Harper, R., Wright, W.: Stories in GeoTime. Information Visualization, 7:3-17. Palgrave Macmillan, 2008

57. Kapler, T., Wright, W.: GeoTime information visualization. In InfoVis '04: Proceedings of the IEEE Symposium on Information Visualization, pages: 25-32, Austin, TX, USA, 2004. IEEE Computer Society

58. Haber, R. B., McNabb, D. A.: Visualization idioms: A conceptual model for scientific visualization systems. In Visualization in Scientific Computing, pages: 74–93. IEEE Computer Society, 1990

59. Granovetter, M.: The strength of weak ties. American Journal of Sociology, 78(6):1360-1380, 1973

60. Renambot, L., Jeong, B., Hur, H., Johnson, A., Leigh, J.: Enabling high resolution collaborative visualization in display rich virtual organizations. Future Generation Computer Systems, 25(2009). Elsevier, 2009

61. ActionTracker. Available from: http://www.princeton.edu/~equids/multimedia.html

62. Kooima, R.: Electro. Available from: http://www.evl.uic.edu/rlk/electro/

VITA

NAME	Mhd Khairi Reda
EDUCATION	M.S., Computer Science, University of Illinois at Chicago, 2009
	B.S., Computer Science, University of Damascus, Damascus, Syria, 2005
EXPERIENCE	Graduate Research Assistant, Electronic Visualization Laboratory, University of Illinois at Chicago, 2008 – 2009
	Intern Software Engineer, Midway Games, Chicago, 2007
HONORS	Honor Graduate, University of Damascus, 2005