

SocialNetSense: Supporting Sensemaking of Social and Structural Features in Networks with Interactive Visualization

Liang Gou*, Xiaolong (Luke) Zhang**
The Pennsylvania State University

Airong Luo†, Patricia F. Anderson‡
University of Michigan

ABSTRACT

Increasingly, social network datasets contain social attribute information about actors and their relationship. Analyzing such network with social attributes requires making sense of not only its structural features, but also the relationship between social features in attributes and network structures. Existing social network analysis tools are usually weak in supporting complex analytical tasks involving both structural and social features, and often overlook users' needs for sensemaking tools that help to gather, synthesize, and organize information of these features. To address these challenges, we propose a sensemaking framework of social-network visual analytics in this paper. This framework considers both bottom-up processes, which are about constructing new understandings based on collected information, and top-down processes, which concern using prior knowledge to guide information collection, in analyzing social networks from both social and structural perspectives. The framework also emphasizes the externalization of sensemaking processes through interactive visualization. Guided by the framework, we develop a system, SocialNetSense, to support the sensemaking in visual analytics of social networks with social attributes. The example of using our system to analyze a scholar collaboration network shows that our approach can help users gain insight into social networks both structurally and socially, and enhance their process awareness in visual analytics.

Keywords: Social network, visualization, sensemaking, visual analytics, SocialNetSense.

Index Terms: E.1 [Data Structures]: Graphs and Networks H.1.2 [User/Machine Systems]: Human Information Processing – Visual Analytics; I.6.9 [Visualization]: Information Visualization.

1 INTRODUCTION

Nowadays, social networks are no longer just about a set of links connecting anonymous nodes, and often contain social attributes of actors and their relationships. Such attribute information as actors' social hierarchy often has significant social implications [6]. In this type of social networks, a person represented as a network node can act both as an individual interacting with other individuals, and as part of a social group connecting with other groups. To make sense complex social networks, we need to consider both network patterns exhibited through the social networks and social structures hidden in the attributes. For instance, in a scholar collaboration network, a scholar has an

affiliation hierarchy consisting of department, university and country, and thus her activities can be investigated at different social levels in the hierarchy, from individual activities to cross-department efforts to international collaborations. Having such information about her collaborations at different social levels could help us better understand how her collaborations have been shaped by her multi-level social affiliations, and how her social status may be influenced by her collaborations.

More importantly, to achieve such understanding, we may also need to go through a series of sensemaking activities, such as exploring and searching collaboration patterns among different social actors and metrics about these collaborations, collecting the social and structural features of interest, organizing them and adding our annotations. These synthesized information forms our understanding of social networks. Furthermore, as the information we gather is accumulated, we might also revise and reconstruct our understanding about the network patterns if we find some conflict information.

However, current methods for social network analysis are often weak in supporting the sensemaking of social networks from both network and social perspectives. Most visual analytics systems only offer analysis of network structures and visualization of these structures, but lack the capabilities to help people interactively explore social networks based on social attributes of a social network. Our recent work, TreeNetViz [15], enables users to exam the hybrid patterns of social and structural features in social networks. However, these tools offer analytical and statistical tools that can produce overwhelming visual outputs, but often overlook users' needs for sensemaking tools that help to gather, synthesize, and organize social, structural and hybrid features of social networks and then construct and revise their mental representations of network data. They also have limited supports for the process of building users' representation, such as where users gather useful information, how users synthesize the information and how they revise their understanding.

In this paper, we present our research work to address these issues. We propose a framework on sensemaking of social networks with an emphasis on integration of network and social attributes to understand social networks structurally and socially. This framework considers both bottom-up processes, which are about constructing understanding of data based on collected information, and top-down processes, which concern using prior knowledge to guide information collection, in analyzing social networks with social attributes, and emphasizes the importance of the externalization of sensemaking processing to visual analytics of social networks. We introduce a system, SocialNetSense, built on the framework to support social network sensemaking activities with interactive visualization, and present a use scenario on how the system helps users understand a scholar collaboration network with social hierarchy information.

2 RELATED WORK

In this section, we review related work from four perspectives: sensemaking theories, visualization for social network sensemaking, visualization of social network with attributes, and

* lug129@ist.psu.edu

** lzhang@ist.psu.edu

† airongl@umich.edu

‡ pfa@umich.edu

process management of visual analytics.

2.1 Sensemaking: Theories and Visualization Support

Sensemaking is a process to successively build and refine representations and fit data with representations to meet the requirements of a task [40]. Generally, sensemaking models fall into two categories: top-down approaches that start with known representations, and then explore data guided by the representations [39][40], and bottom-up approaches that first explore data and then construct presentations of data based on relevant information collected [4]. These two approaches often work together in the process of sensemaking, and therefore some hybrid models have been proposed to integrate them [26] [37].

Visualization has been used to support various sensemaking activities. Some visualization systems are proposed to support general sensemaking. Chi et al. [9] visualize a spreadsheet to help the sensemaking of the evolution of web sites. Jigsaw [44] uses visualization tools to identify key information entities and their relationships in support of the sensemaking of a large collection of text documents. CzSaw [24] supports iterative analysis of document collection with editable and re-playable history navigation. Sandbox [46] enables users to search, collect, organize and marshal evidences to make sense documents in a visual space. To help scholars to make sense of literature, CiteSense [48] provides an integrated visualization environment to support literature search, selection, organization and comprehension.

These visualization systems show some features that are important to sensemaking activities. These features include searching, collecting, organizing relevant information, linking concepts, creating and modifying representations, and tracking the sensemaking progress. However, they were designed to support sensemaking of general types of data, such as documents and web sites, and cannot be applied for social network sensemaking.

2.2 Social Network Sensemaking with Visualization

Sensemaking activities of social networks involve network exploration and representation creation. Visualization tools are designed to support these two types of sensemaking activities.

To support network exploration, most social network visualization tools support network analysis, visualization and interactions. These tools can be classified into two categories: toolkit-oriented and application-oriented. Toolkit-oriented designs mainly focus on the general layout of networks but their support for social network analysis and user interaction is usually limited. Tools like Prefuse [18] and JUNG [34] provide different layout algorithms, and serve as libraries to build customized graph visualization tools. Some tools, such as Pajek [1] and Visone [5], are designed for researchers to conduct traditional social network analysis with rich analysis functions, such as identifying important actors and actor roles, but their support for visualization and user interaction is weak. Application-oriented tools like Vizster [17], MatrixExplorer [21], NodTrix [22], and NetLens [25] support user exploration of and interaction with social networks, but the exploration tasks they support are very specific, not broad enough for complex sensemaking activities.

Sensemaking of social networks involves helping users build representations with an iterative process in both bottom-up and top-down fashions. Some visualization systems partially support this process. For example, SocialAction [35] allows iteratively ranking and filtering nodes; SYF [36] can guide domain experts to explore network data; and Shrinivasan and van Wijk [43] designed a system to support the analytic process for general information visualization. However, few visualization tools integrate both top-down and bottom-up approaches to support sensemaking of social networks.

In sum, most tools mentioned above are often weak in

supporting the sensemaking of social networks from both network and social perspectives. Most studies emphasize network structures, or only employ attributes to filter, sort, or group actors. Few can help users analyze the impacts of social attributes on structural patterns of social networks, such as examining the network features (e.g., size, degree, importance, etc) from social perspectives. Also, these systems only offer a number of analysis and statistical tools that can produce overwhelming visual outputs, but cannot help user synthesize related analysis results to build up representations of networks. Few can support the management of analysis process.

2.3 Visualization of Social Network with Attributes

Visualization of social networks with social attributes often requires revealing network patterns over social attributes, such as identifying important actors with a specific attribute, their connections, and relevant social attributes (e.g. affiliation). Some designs used multivariate graph visualization to present a graph with attribute data of nodes or edges, and some leveraged hierarchical social attributes to examine network patterns at different levels.

Multivariate graph visualization largely focuses on network clustering and aggregation based on the attributes. Some techniques emphasize the layout of the graph based on the multivariate data, such as multivariate state transition graphs [38] and other techniques consider user interaction and navigation, including PivotGraph [30], which aggregates and places nodes on a grid of two dimensions of attributes specified by users; Semantic substrates [42], which organizes network nodes based on user-defined regions of attributes, and GraphDice [3], which uses a plot matrix to present node-link plots for every combination of attributes of actors and edges in a social network. Other tools abstract data attributes with network schemas to help users to explore data. For example, OntoVis [41] visualizes networks with an ontology-based schema and Ploceus [29] enables users to construct network semantics of multivariate tabular data and conduct multi-dimensional and multi-level analysis with semantics. However, in these systems, attribute information used in building hierarchy is either hidden in network structures or not explicitly available for users to interact and explore.

Some methods have been proposed to visualize network patterns over hierarchical social attributes by integrating both network and tree. Several approaches directly add network connections layered over treemaps [12] [32]. Along this direction, some efforts have been made to reduce edge crossings or node occlusion [7] and to avoid edge cluttering [10]. However, these designs largely focus on representing networks with attributes at a single social level. Recently, researchers have explored designs to combine both social attributes and networks in support of user interaction with and exploration of network patterns at different levels of social aggregation [15] [45].

2.4 Process Management of Visual Analytics

Process management is an important component in visual analytics. It focuses on users' reasoning process in interacting with various visualizations. It captures, records and regenerates user interactions and helps users keep track of the development of their representations about data. The work of process management of visual analytics can be organized around two tasks: visualization exploration and representation creation.

Most research of process management of visual analytics focuses on visualization exploration. It captures users' interactions with visualization and stores them with different models to regenerate visualization results. Tools falling in this category usually provide simply undo and redo operations to navigate a

series of interaction history [31]. However, a linear model of redo and undo is insufficient to manage complex exploration process. Researchers proposed some models of process management based on state and action of visualization [19]. Kreuseler et al. [27] use dependences of state transitions and a history tree manage a visual mining process. Jankun-Kelly et al. [23] propose a model to capture, share and reuse user interactions in visualization exploration with a parameter set, and transformation of the parameter set. Heer et al. [19] design graphical history tools for Tableau to record and visualize interaction histories. Some designs capture transitions of visualization states with scripts of actions. VisTrails [2] uses a script-like visual language to record and execute pipelines of scientific visualization. Czsaw [24] uses a similar approach to record a document analysis process.

However, few studies provide the process management of users' representation generation process. While some works use various semantic annotations to history of interactions with visualization, such as tagging [43], bookmarking [19][20] and text annotating [16], supporting the history manage of users' understanding generation, such as collecting and organizing evidence, composing and modifying arguments, creating links among different elements is still limited. Part of our work reported in this paper is about a process management model of users' representation building in social network sensemaking.

3 A FRAMEWORK OF SOCIAL NETWORK SENSEMAKING

We first present a sensemaking framework of social network with social attributes as shown in Figure 1. It involves three parts of social network data, network exploring loop and representation building loop. In this framework, social network data with attribute information are first processed and reorganized into different structures and visual representations with emphasize on social features, network features and hybrid features. The data processing, including data modeling and network aggregating, is discussed in section 3.1 and section 3.2.

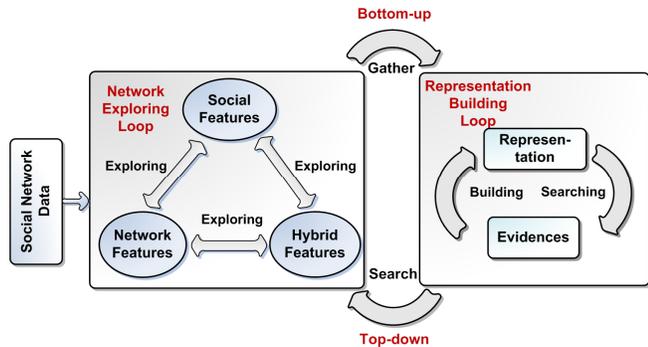


Figure 1: A framework of social network sensemaking with visualization.

Sensemaking activities go through two loops in our framework: a network exploring loop and a representation building loop. In the network exploring loop, users seek desirable information by exploring three features of network data, including social features, network features and hybrid features, according to a task and existing knowledge. The social features are derived social attributes of actors in a network, such as hierarchical affiliation information of department, college and university. Network features emphasize on structural information, such as importance of nodes and links, network cliques, and other network properties. Hybrid features show the impact of social attributes over network structure, and focus on the aggregated network patterns among actors from different attributes, such as the collaboration patterns among one university with other departments in a scholar

collaboration network. The three features are related and various metrics (such as degree, betweenness and closeness), plots (such as degree distribution) and visualization tools can help users to explore these features.

The two loops can interact with each other and form an outer loop of bottom-up and top-down processes. In the top-down process, users can use the representations they have to guide their further search in the exploring loop to look for new evidences. In the bottom-up process, users actually collect information of interest as evidences to support or revise their representation. The evidences are gathered and used in the representation building loop. As the evidences are accumulated, representations can be modified and even re-constructed.

The sensemaking process of two inner loops with the outer loop indicates the trajectory of moving from a state of non-sense to a state of sense. Operations in both network exploration and representation building loops are important to help users to understand their sensemaking process. The intermediate analysis results are important to understand the process because users often revisit an intermediate stage during the sensemaking process. These states should be stored in an automatical fashion or upon a user's demand. Thus, a process management is usually provided in the sensemaking activity to enhance the awareness of their sensemaking process and help users track their analysis trajectory. The process management mechanism is discussed in Section 3.3.

The framework captures key analysis behaviors and process of users to make sense of social network with attributes. In the following section, we use the framework to guide our design of data model of social network with attributes, multiscale and cross-scale aggregation, process management and functions of our visualization system.

3.1 Data Model of Social Networks with Attributes

In this section, we introduce a graph model to represent a social network with attributes. The social attributes may include information with different social entities, such as department, university in a scientific collaboration network. In this sense, this type of network is a special case of a multimodal graph, in which the nodes and edges have different modes, or types [13]. Here we define the network of interest as follows.

Suppose we have a social network, $N = (V_N, E_N)$, where V_N are social actors (nodes) and E_N are social relations (edges), and each node has m attributes. Some attributes may have hierarchical levels. For example, the social affiliation attribute of an actor in a research collaboration network can be described at the levels of individual, university, and country. Therefore, we can denote a hierarchical attribute as a tree structure, $T_i = (V_{T_i}, E_{T_i})$, where V_{T_i} are tree nodes derived from the i th attribute and E_{T_i} are tree edges indicating parent-child relations. Note that even if an attribute only have one level of element, we can construct a two-level hierarchy by introducing a dummy root node.

To analyze a social network with social hierarchical attributes, we use a **TreeNet** graph model [15], a compound graph consisting of two sub-graphs: one tree structure and one network. A TreeNet graph, $TrNi$, can be written as:

$$TrNi = (Ti, N, \rho),$$

where ρ is a function to map tree leaf nodes to network nodes: $\rho(n_{Ti}) \rightarrow n_N, n_{Ti} \in V_{TiL}, n_N \in V_N$, and V_{TiL} is the set of leaf nodes in the tree Ti . The TreeNet graph model is used to aggregate network at different levels of attribute, as discussed in next section.

3.2 Multiscale and Cross-scale Aggregation of Social Networks with Attributes

A TreeNet compound graph can help the understanding of the hybrid features of social and structural patterns in social networks.

In this case, we can have both *multiscale* and *cross-scale* relationships among social entities. The concepts of multiscale and cross-scale relationships are similar to within-mode and between-mode edges in a multimodal graph context. Multiscale relationships are the ones between entities from the same level, such as between individual and individual, or between university and university. Cross-scale relationships refer to those links between entities from different levels, such as between individual and university, between individual and country, or between university and country. The detailed algorithms of multiscale and cross-scale aggregation can be found in [15].

An example of network aggregation is shown in Figure 2. Suppose we have a co-authoring network shown in Figure 2a, in which each node is an author and there is a link between two authors if they co-author at least one paper. Each author has an affiliation attribute of department and college, which can be represented as a tree shown in Figure 2b (we assume each author only has a single affiliation). For example, author 0 is from department *D* in college *A*. Thus, if we are only interested in the nodes specified by a cutline *CL* shown in Figure 2d, we can collapse nodes *A* and *D*, and expand node *E* and *F* to show the author nodes under them. In this way, we actually aggregate the nodes in college *A* and department *D* and also aggregate the edges of authors under *A* and *D* but connecting with other nodes. For example, two edges $e(0, 3)$ and $e(1, 3)$ are aggregated into an edge $e(A, 3)$. Therefore, we can have an aggregated network view shown in Figure 2c which spans three levels of college, department and individual author.

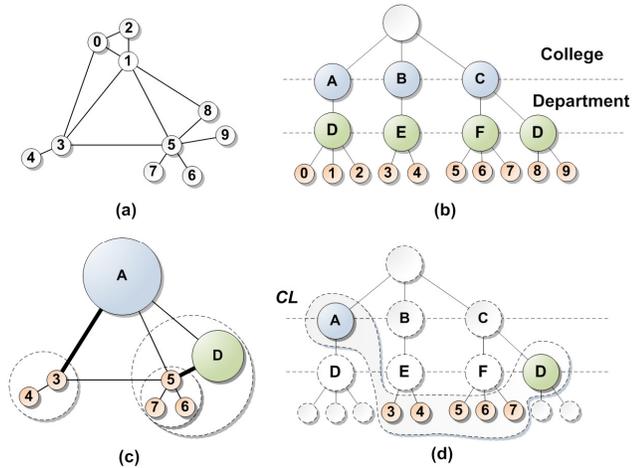


Figure 2: A Network aggregation example in a TreeNet Graph.

3.3 Process Management of Sensemaking with Visualization

Process management of sensemaking helps users keep track of their reasoning process with interactions of visualization. It involves interactions of both visualization exploration and representation creation. In this paper, we emphasize on users' representation building process. We argue that interactions of visualization exploration are too complicated and these interactions alone are not sufficient to help users to make sense of the dataset. It is more important to capture the process of representation developing, because users should know how and from where they develop their current understanding of data of interest. Furthermore, study shows that relationships in data can be recoverable through manual inspection [11]. Users can revisit an intermediate representation and manually regenerate

visualizations with records of the intermediate representation. Thus, our process management mechanism is mainly targeted to supporting users' representation building process.

We use a hybrid model with state and action of visualization by following the history management paradigm suggested by Heer [19]. Our hybrid model (shown in Figure 3) uses a tree structure to manage states and actions. It records all user actions to build up their representations, but only caches critical states of the representation developing. A state, S_i , is a stored configuration of a user's current representation of visualizations, charts, notes, groups and so on. A critical state is either an initial state ($Sc0$ in Figure 3) or a branching state ($Sc2$ in Figure 3). A state becomes a branching state when a user returns to this state and this state is transformed to a new state by another action. A cached branching state is useful to reduce the number of actions to be reversed when a user wants to jump to a specific state. For example, if the system is currently in state $S21$ and a user wants to go to $S22$, the model only needs to retrieve the cached branching state $Sc2$ and execute action $A4$ to achieve this.

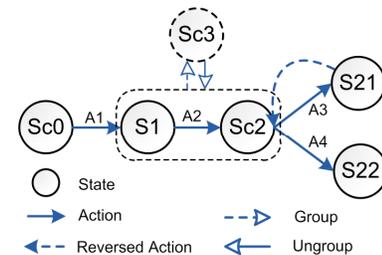


Figure 3: A hybrid history management model of states (S_i) and actions (A_j). It only caches the initial state ($Sc0$) and a branching state ($Sc2$). $Sc2$ is a branching state because the system is reversed from state $S21$ and is transformed into state $S22$ by a new action $A4$. States $S1$ and $Sc2$ can be grouped into state $Sc3$, which also can be ungrouped.

Table 1. Actions included in our hybrid history management model.

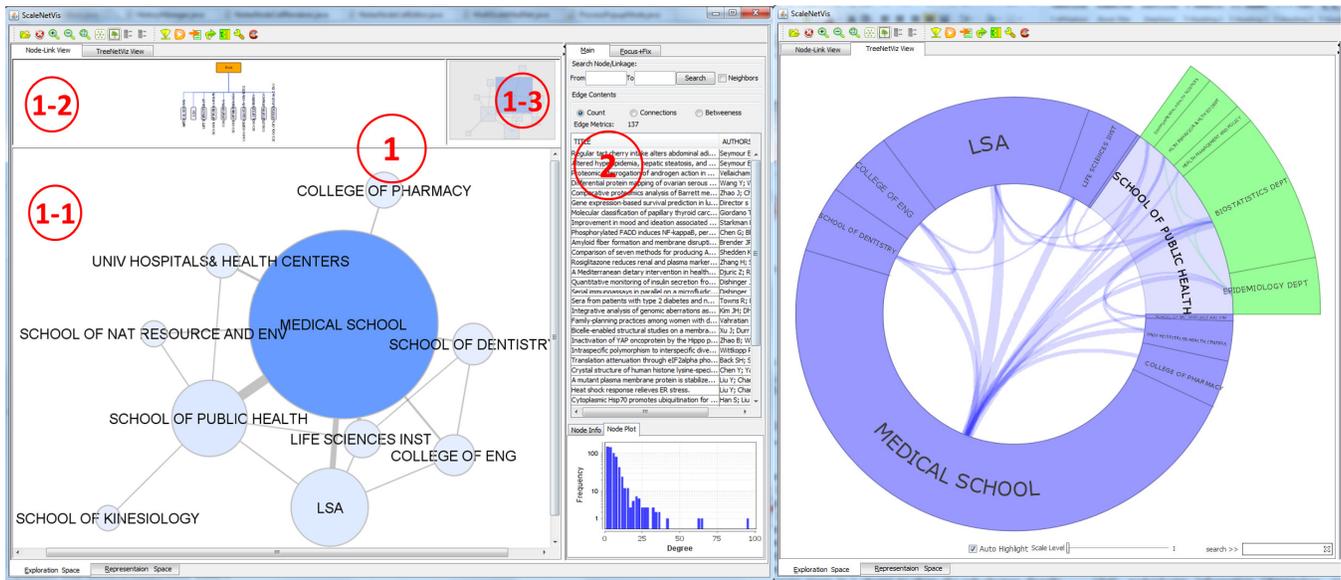
| Visual Insight Actions | Knowledge Insight Actions | Meta Actions |
|--------------------------|---------------------------|----------------|
| Gather Visualization | Create/Edit/Remove Note | Resize Move |
| Gather Plot | Link/Unlink Elements | |
| Gather Node/Edge Metrics | Group/Ungroup Elements | |

An action, A_j , is a meaningful and atomic operation by a user to create their understanding of visualization. Each action connects two states and is a transition among two states. The actions used in our model are shown in Table 1 by following a taxonomy schema proposed by Gotz and Zhou [14]. A series of actions can be easily merged and unmerged in the hybrid model. In Figure 3, the two states $S1$ and $Sc2$ are grouped into a new branching state $Sc3$, so that two actions $A1$ and $A2$ are merged. This is helpful when users want to simplify process by merging multiple actions into a single meaningful action when a history tree becomes complicated.

4 SOCIALNETSENSE: A VISUALIZATION SYSTEM TO SUPPORT SOCIAL NETWORK SENSEMAKING

Guided by the sensemaking framework of social-network visual analytics, a new visualization system should support:

- *social features exploration* to help users to understand social structures embodied in attributes of actors, such as a hierarchical organization structure of an individual, the peers sharing similar social status of an individual.



(a)

(b)

Figure 4: Network Exploring Space (NES) in SocialNetSense: (a) Coordinated node-link views, including a network visualization (Panel 1-1), a tree visualization (Panel 1-2) to show the social hierarchy of actors, a network overview panel (Panel 1-3) to support quick navigation, and a control panel (Panel 2) with a set of analytical tools of different aggregation metrics, analytic plots and so on; (b) A TreeNetViz view to reveal hybrid features of network patterns over a social hierarchy.

- *network features exploration and analysis* to help users to find important and peripheral nodes, critical edges and closely connected clusters, and exploration should be supported with different visual representations so that users can triangulate their understanding of network patterns with different views;
- *hybrid features exploration* of a network over a social attribute to help users understand aggregated network patterns at the same and different levels of a social attribute.
- *representation building-up* to allow users to collect visualizations of subgraphs, plots, metrics, group them, add notes to create and modify their understanding of data;
- *process management* to enhance users' awareness of analytic process by capturing, storing users' interactions which form their representations, visualizing their action history, and enabling users to return to a specific intermediate representation.

With these requirements, we design and implement SocialNetSense to support sensemaking activities of social network with attributes from both social and structural perspectives. In this section, we present the visualization and interaction design of SocialNetSense, including three main components: a Network Exploring Space (NES), a Representation Building Space (RBS) and a process view.

4.1 Network Exploring Space (NES)

The interface of Network Exploring Space (NES) is shown in Figure 4a. It includes two main panels: a visualization view (Panel 1) showing social networks along with hierarchical social structures, and a control panel (Panel 2) offering a set of analytical tools, such as different aggregation metrics, analytic plots and searching. The tool bar above Panel 1 includes view manipulation tools, such as zooming, panning and layouting.

Our tools support multiple forms of visual representations of social networks to help users to explore social, network and hybrid features. A traditional node-link style to view a network, as shown in Panel 1 in Figure 4a, is provided to present social

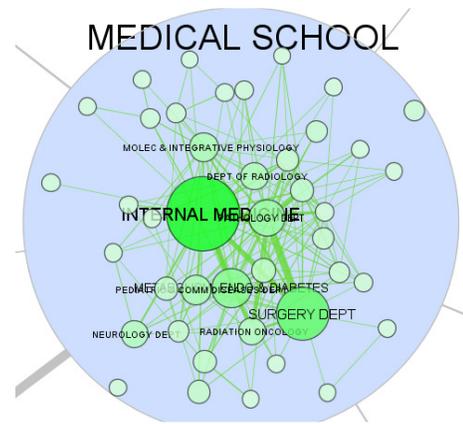


Figure 5: The collaboration patterns of department under Medical School after semantic zooming in a node-link view.

features and network features of the data, and a TreeNetViz view (Figure 4b) to see the hybrid features of how a social network is coupled with social structures of its actors. Different views offer users insight into network and social features from different perspectives. The node-link network view is intuitive and easy to identify highly-connected nodes, but requires extra cognitive efforts to connect network features with a social hierarchy. TreeNetViz view integrates network and the tree structure in a single view and offers easy aggregation and dis-aggregation operations to understand social ties at different levels of social aggregation.

4.1.1 Coordinated Node-link Views

The node-link view includes three view panels, shown in the panel 1 of Figure 4a: a network view panel as the primary workspace for network visualization (Panel 1-1), a tree view panel (Panel 1-2) to show the social hierarchical features of actors, and a network overview panel (Panel 1-3) to support quick navigation.

General zooming and panning are provided for the tree view

and network view. The two panels are also resizable to offer user more flexibility to exam both network and tree structures. Furthermore, semantic zooming is also provided to explore a network structure at different social scale. Users can also specify a set of nodes as focused nodes which are affected by semantic zooming while other nodes are fixed to their current scales. Users can select them by clicking those nodes with the CTRL key down. This generates a cross-scale view. For example, in the Panel 1-1 of Figure 4a, we can see the overall network patterns of collaboration at the level of school. After semantic zooming in the node of Medical School, we can have a view shown in Figure 5, which indicates the collaboration patterns of departments under the node of Medical School in details.

Three view panels facilitate navigation and interaction with hierarchical aggregated networks. The primary network view and the tree view are coupled together to support user navigation. When users zoom in a non-leaf node to show the child nodes directly under it, the tree view is simultaneously expanded to display its child nodes. Similarly, when zooming out to a more abstract scale, users see the tree view collapses to a level at which all the child nodes are hidden under their parent node. In addition, node selection is synchronized in two views. When users select a node in the network view, the corresponding node in the tree view is also highlighted.

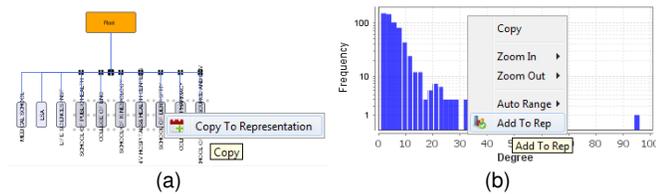


Figure 6: Two examples of evidence collection widgets in NES: (a) collecting a subtree; (b) collecting a degree distribution plot.

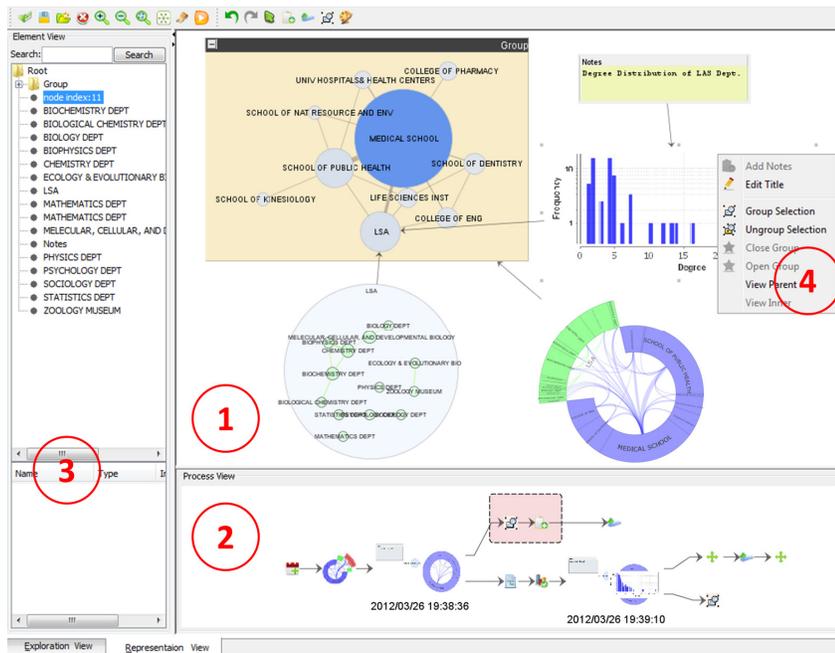


Figure 7: Representation Building Space (RBS): Panel 1 is the main working space; Panel 2 is the process view; Panel 3 lists the elements in the working space; and Toolbar and Popup menu provide tools of representation building, such as add note, group/ungroup elements and link element.

The overview in Panel 1-3 of Figure 4a uses a rectangle to represent the location and size of the current network view. When users manipulate the network view through panning or zooming, the location and size of the rectangle will be updated automatically. Users can directly resize or change the location of the rectangle in the overview to achieve quick navigation in the network view.

4.1.2 TreeNetViz View

We include a novel visualization technique, TreeNetViz [15] (shown in Figure 4b), to reveal hybrid features of network patterns over a social hierarchy. This visualization technique supports interactive aggregation of a network based on the attributes of a tree structure, and allows users to explore multiscale and cross-scale relationships among different social entities (e.g., individuals, groups, and organizations) in a single view. TreeNetViz uses a Radial, Space-Filling (RSF) method to represent a tree structure, a circle layout show a network and edge bundling to avoid visual cluttering. Besides interactions of multiscale and cross-scale exploration, it also supports ego-network view to show direct neighbors and their links. For example, in Figure 4b, the nodes of schools, such as Medical School and School of Public Health are assigned along a circle. The child nodes are assigned to the sub-arc of their parent node. We can see there are some department nodes under the School of Public Health. In this way, we see the network patterns among the departments under School of Public Health with other schools. Users can expand and collapse node sectors by double clicking, or resize the size of node sectors with a mouse wheel. The details of this visualization can be found in [15].

4.1.3 Evidence Collection in NES

In NES, we offer users handy tools to collect evidence of interest to build up their own understanding during their exploration. One important task in the bottom-up process of sensemaking is to search information of interest during exploration and collect them

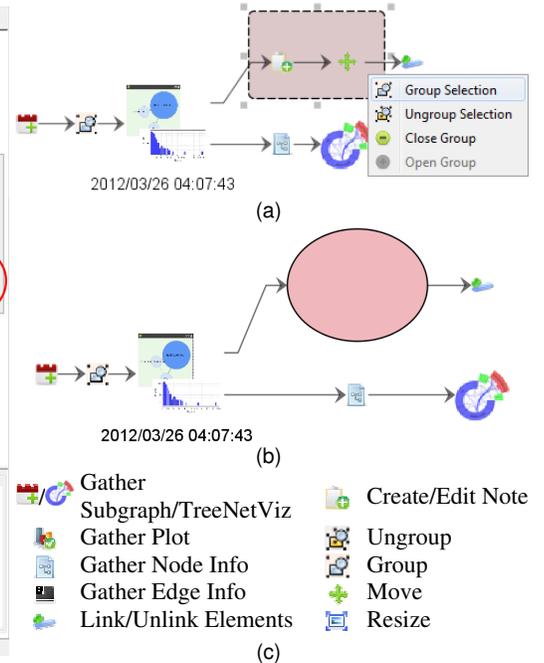


Figure 8: View of process management: (a) a popup menu to group two actions of "add note" and "move"; (b) the grouping result of (a); (c) icons of major actions used in the process view.

as evidences to support or revise their representations. In NES, there are various collectable sources of information, including node-link views of sub-tree and sub-network, a TreeNetView of network, plots of various metrics (e.g. a degree distribution plot), node/cluster information (description, metrics of size, degree, betweenness, closeness) and edge information (description such as paper information in a co-authoring social network, metrics of betweenness and weight). Users can right click on information source of interest to popup a collecting widget and then copy it into Representation Building Space (RBS). Figure 6 shows two examples of evidence collection widgets in NES. Figure 6a shows that using a pop-up menu to copy a sub-tree into RBS and Figure 6b shows copying a degree distribution plot into RBS.

4.2 Representation Building Space (RBS)

A Representation Building Space (RBS) allows users to organize the evidences collected from the NES, establish relations among them, take notes and create representations. Figure 7 shows the user interface of RBS. The interface includes three main panels: a working space (Panel 1), a process view (Panel 2) and an element list view (Panel 3). In the working space, users can save visual structures (such as visualizations from the node-link view and the TreeNetViz view), network metrics (such as size, degree, betweenness and closeness) and plots from the NES into this space, organize (group/ungroup) them, add notes, create connections among them and build representation in a bottom-top way. Also, users can use generated representations to guide evidence collections in the NES, such as deciding which tree nodes should be visible to generate a new aggregated view on a social network. This is a top-down process.

This space can be manipulated by general edit and view control tools like redo, undo, zooming, panning, and so on. Annotation tools are provided to help users to form their understanding of the collected network information. Notes, connections and groups can be added on visual structures in the working space by various tools (Toolbar and Popup menu 4 in Figure 7). Panel 3 lists the elements in the working space, including tree hierarchies, aggregated networks, statistics plots, notes and connections. Users can conduct a search of these elements in Panel 3.

4.3 Process View

A process view (Panel 2 in Figure 7 and Figure 8) realizes the process management model discussed in Section 3.3. Evidence collection and representation building-up actions (shown in Table 1) are captured, recorded and visualized in this view. This view visualizes an overview of sensemaking activity history with a left-to-right tree layout (Figure 8a and 8b). The visualization is a little different from the model shown in Figure 3: nodes indicate actions and links means the sequences of actions. The concept of state is implicit in this visualization. We use snapshots of representation space with timestamp at branching nodes where users have different paths of representation building actions. Thus a snapshot can provide an anchor point to remind users the intermediate representation of analysis results in navigating the whole analytic process. A selected node indicates the state at which users' current representation view is. We use different icons to represent actions shown in Table 1, and the icons' meaning is shown in Figure 8c. When users hover mouse over an action node, it also shows details of its action type and the time when it was executed.

We provide handy interactions in this view to help users navigate in a process of representation building. First, the process view is zoomable: users can zoom into to check details of operations and visual structures, or zoom out to see the whole trace of sensemaking activities. Furthermore, users can *revisit* any

intermediate representation in an analysis process by double clicking a tree node which is not the current state. The system will locate the nearest branching node and retrieve the cached state of that branching node, and then execute the actions from the branching node to the selected node. In addition, users can select a series of actions and use a popup widget to group them. The grouped actions can be also ungroup. For example, Figure 8 (a) shows the popup menu to group two actions of "add note" and "move", and Figure 8 (b) shows the grouped result.

5 USE SCENARIO: MAKE SENSE OF A CO-AUTHORING SOCIAL NETWORK OF MDLINE WITH SOCIALNETSENSE

In this section, we present a scenario using SocialNetSense to analyze a co-author network of diabetes researchers at University M and we want to use this system to help us make sense the collaboration patterns among these researchers. We select diabetes researchers as an example because they are from diverse background including chemistry engineering, public health, life science and biomedical research.

The dataset includes 847 authors and 2,498 co-author relationships on 614 papers (two authors are considered to have a link if they co-author on at least one paper). In this example, we focus on the influence of affiliation attribute of authors over a network structure and thus we use two levels of affiliation: college and department. In the dataset, we identified 10 college-level nodes and 90 department-level nodes. A school and a research centre are also treated as college-level actors. Detailed information of the dataset can be found in [15].

This use scenario aims at understanding how SocialNetSense supports users' sensemaking of the patterns of co-authoring collaboration across organizational boundaries. Specifically, we are interested in sensemaking activities to crystalize users' understanding about the social network with affiliation attributes occurring in SocialNetSense, such as:

- How do different visualizations and metrics help users form their understanding about collaboration patterns vary across departments and colleges?
- How are users supported to identify those researchers acting as "boundary liaison" to connect different departments and colleges?
- How does the process view support users to revise their understanding during the sensemaking process?

5.1 Sensemaking of Multiscale and Cross-scale Network Patterns

SocialNetSense can help a user to build her understanding of the collaboration patterns of the diabetes researcher network, such as the power and status of social actors and collaborations, at three different levels of colleges, department and individuals. Figure 8 shows an example of a representation of the main network pattern at the level of college composed by a user in the NES.

Let us take a look at how the user got this representation. First, the user explore the network in NES with the node-link diagram view found there are three dominant colleges of "Medical School", "LSA" (Literature, Science and the Arts) and "Public Health" at the level of college (shown in Panel 1-1 in Figure 4a). Node size represents the number of researchers in a group and edge thickness shows the collaboration strength between two groups. The user believed this pattern is important and thus copies the subgraph of three colleges into RBS with an evidence collection tool shown in Figure 6a and grouped them with a popup tool shown in Figure 7-4. This group is shown in Figure 9-1. Then, the user also collected metrics of degree for the three nodes and the metrics were saved as text boxes (2a, 2b and 2c in Figure 9). The users also linked the metric information to the node by

dragging the text box to the related node. A degree distribution plot was also saved with a tool shown in Figure 6b and linked to “Medical School” as shown in Figure 9-3. Figure 9-4 shows the gathered information about the papers which are co-authored between “LSA” and “Public Health”. All information was grouped with a note shown in Figure 9-5.

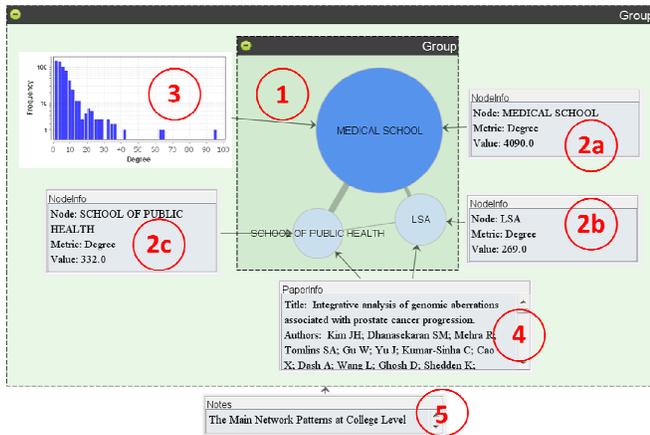


Figure 9: A user’s representation of the main network patterns at the level of college.

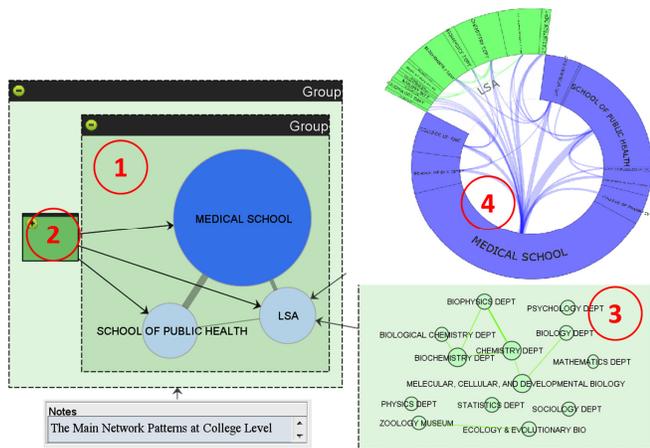


Figure 10: A user’s representation of both the main network patterns at the level of college and cross-scale patterns among the department in LSA with other colleges.

SocialNetSense also supports a user to make sense cross-scale network patterns. By referring to the representation in Figure 9, the user went back to NES and continued gathering more information to understand cross-scale patterns among the departments in “LSA” with other colleges. To simplify the representation, the user first grouped node and edge information into a folder shown in Figure 10-2. Then, in the node-link diagram view, she used semantic zoom and get the details of the network patterns under college of “LSA” shown in Figure 10-3. The user collected these patterns, saved and grouped them in RBS. To show the cross-scale network patterns among “LSA” and other colleges, the user went back to NES and opened a TreeNetViz view. She also gathered a TreeNetViz view with an expanded sector of “LSA” shown as Figure 10-4 into the representation. Finally, as shown in Figure 10 the user built her understanding of both the main network patterns at the level of college and cross-scale patterns among the departments in LSA with other colleges. The user could continue revising this representation according to

her tasks.

5.2 Sensemaking of “Boundary Liaison”

SocialNetSense also supports a user to make sense of the network with affiliation to identify important actors acting as “boundary liaison” to connect different departments or colleges. When a user examined the network patterns at college level (shown in Panel 1-1 in Figure 4a), she found that “Medical School (MS)” is not directly connected with “School of Kinesiology (SK)” and they are linked via “School of Public Health (SPH)”. She gathered the subgraph of the three nodes in to RBS and takes a note of path “MS--SPH--SK” (Figure 11-1). Then, she wondered which department under “SPH” collects the two schools of “MS” and “SK”. Thus, she went back NES and used a TreeNetViz view to expand SPH to the level of department, and found it is “Hlth Behaviour & Hlth ED Dept (HBHED)” connecting MS and SK at the department level. She gathered and saved this information with a note in RBS (Figure 11-2). Similarly, she finally identified the “Auth525” is the boundary liaison connecting the two schools of “MS” and “SK” (Figure 11-3). Using SocialNetSense, the user created her understanding of the roles of boundary liaison at different levels. With this representation, she can further examine network metrics of these critical actors.

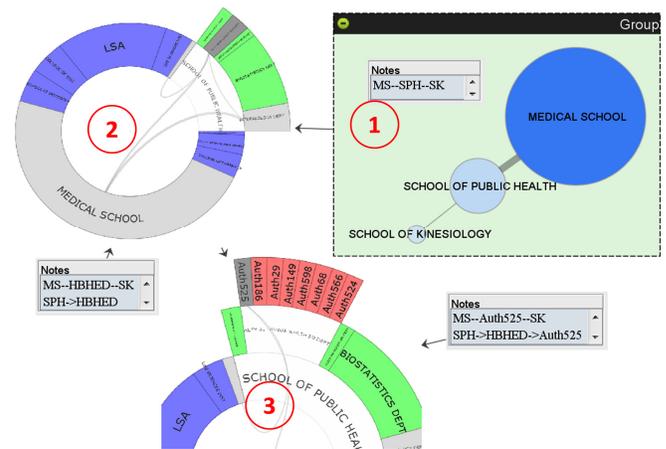


Figure 11: A representation of boundary liaisons connecting “Medical School (MS)” and “School of Public Health (SPH)”.

5.3 Representation Revising with Process View

The process view of SocialNetSense visualizes users’ interactions with representation space and helps users understand the representation building process. It also helps users to navigate in a process and revisit intermediate analysis results. Figure 12 shows the representation building process of the example in Figure 11. We can see that the user did not follow a linear analysis process to get the final representation. After the user found the path of “MS--SPH--SK” (Figure 11-1), there is an important change of analysis strategy which is shown at the branching node of Figure 12-1. With a close look at this branching node, we find that the user first tried to use a node-link view to examine which department in “SPH” connects “MS” and “SK”. This is shown in the lower subtree of the process path in Figure 12. The action in Figure 12-2 is gathering a subgraph with a node-link view into RBS. The subgraph, shown in Figure 12-3, was generated by using semantic zooming with the focused node of “SPH”. We can see the collaboration details under the node of “SPH” in this subgraph, but it is not very helpful for the user to identify the boundary liaison in the school of “SPH”.

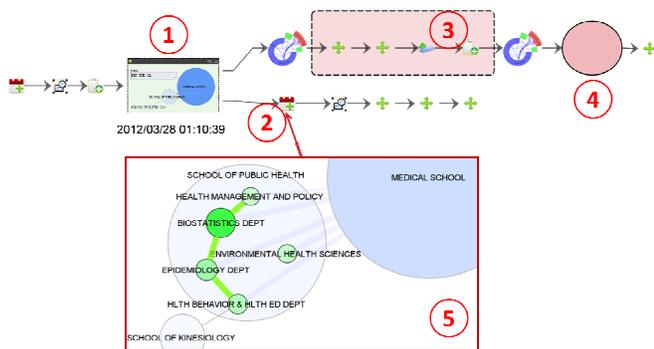


Figure 12: The process view of the representation building-up process in Figure 11.

Therefore, the user abandoned the strategy of node-link view by returning to the state of branching node (Figure 12-1) and started to use a TreeNetViz view to search the liaison actor between “MS” and “SK” in the school of “SPH”. The new strategy is shown in the upper subtree of the process in Figure 12, which helps her to reach the final representation. We can see icons of TreeNetViz view in this branch of analysis process.

We can also see that some actions are grouped together here to simplify the analysis process in this view. For example, Figure 12-3 shows an expanded action group and Figure 12-4 shows a collapsed action group. The user can easily return to any intermediate analysis result in two different paths of strategy by double clicking the action nodes in this view.

The use scenario shows that SocialNetSense allows users to make sense of social network from both structural and social perspectives. With SocialNetSense, users can examine collaboration patterns of social entities at multiple and cross levels. It provides users sensemaking tools to gather, synthesize, and organize various visualization and metrics to create their own understanding about network data. The process view enables users to examine the whole analysis process and revisit intermediate analysis results. We can see that SocialNetSense provides a start point to have a general understanding of the dataset. With this understanding, researchers can further answer some complicated questions such as how the main patterns of network are formed, what resources the boundary liaisons have in collaborative projects, and how we can help research seek potential collaborators.

6 DISCUSSION AND CONCLUSION

In this paper, we presented a framework for the sensemaking of social networks with social attributes from both social, structure and hybrid perspectives. This framework is built upon sensemaking theories and integrates both data-driven and representation-driven approaches, and emphasizes interactions between them. We introduced a prototype system, SocialNetSense, to support a two-way process of social network sensemaking activities, including exploring and examining network patterns with node-link views and a TreeNetViz view, and gathering various visualizations, metrics information and plots, organizing them, adding notes to create user representation of social network with attributes. We also presented a use scenario utilizing our system to make sense of a co-authoring social network and show its potentials to help users crystalize their understanding of social networks from both structural and social perspectives.

The contributions of this work fall into two categories: theoretical and practical implications for visual analytics of social networks. On one hand, we propose a sensemaking framework for

visual analytics of social networks with attributes. This framework integrating both top-down and bottom-up processes can be applied to visual analytics of general social networks. Two aspects with the sensemaking approach to general social network visual analytics are: 1) to enable users to explore social networks from different perspectives and 2) to allow users to collect, organize and synthesize information of interest to build their understanding. On the other hand, we present a visual analytic system for social network. With using the system, we can validate and revise our sensemaking framework, and then derive design implications for other visual analytic systems of social networks.

Another interesting issue in this work is generalization of the design of Representation Building Space. We already see some similar design for different sensemaking tasks, such as CzSaw [24], Sandbox [46] and CiteSense [48]. Theoretically, we can have some common sensemaking activities which should be supported to help users to build their understanding, such as collecting, organizing, synthesizing and annotating information. Practically, it is difficult to have a general technical mechanism to support this representation building because of the speciality of domain tasks. However, we might develop a technical mechanism for visual analytics of social networks by incorporating commonly used visualization of network, such as node-link view, matrix view and circular view.

The research has some limits. First, our approach may not perform well in dealing with data with incomplete attribute data. If the attributes of some key nodes are missing, it becomes difficult to construct aggregated networks and to support exploring social duality. Secondly, the process management module currently only captures, stores and executes the interactions occurring in representation spaces. Although we argue that it is more important to capture the process of representation development, history of interactions during visualization exploration are also helpful to automatically regenerate visualizations of interest.

We will extend our work in three directions. First, we will develop methods to help create social networks by integrating various information sources (e.g., organization directories, research paper repositories). Secondly, we will extend our process management module to capture interactions during visualization exploration. Thus, our system can automatically regenerate visualizations used in users’ representation. We also plan to provide handy tools to import and export the representation users built during the sensemaking process. Finally, we will deploy the system into fields and work with domain experts to collect data about how they use our system to support their sensemaking activities. We can analyze the usage data to evaluate our sensemaking framework and improve the system.

REFERENCES

- [1] Batagelj, V., & Mrvar, A. (2005). *PAJEK: Program for Analysis and Visualization of Large Networks*. Ljubljana, Slovenia: University of Ljubljana.
- [2] Bavoil, L., Callahan, S., Crossno, P., Freire, J., Scheidegger, C., Silva, C., & Vo, H. (2005). Vistrails: Enabling Interactive Multiple-View Visualizations. In *Proc. of InfoVis’05*, pp. 135-142.
- [3] Bezerianos, A., Chevalier, F., Dragicevic, P., Elmqvist, N. & Fekete, J. D. (2010). GraphDice: A System for Exploring Multivariate Social Networks. *Computer Graphics Forum*, 29(3), pp. 863-872.
- [4] Bodnar, J. (2005). Making Sense of Massive Data by Hypothesis Testing. *International Conference on Intelligence Analysis*, pp. 2-4.
- [5] Brandes, U., & Wagner, D. (2003). *Visone: Analysis and Visualization of Social Networks*. *Graph Drawing Software*. Springer, Berlin, pp. 321-340.

- [6] Breiger, R. (1974). The Duality of Persons and Groups. *Social Forces*, 53, pp. 181-190.
- [7] Burch, M. & Diehl, S. (2008), TimeRadarTrees: Visualizing Dynamic Compound Digraphs. *Computer Graphics Forum*, 27, pp. 823–830.
- [8] Card, S.K., J. D. Mackinlay, & B. Shneiderman, (Eds.) *Readings in Information Visualization: Using Vision to Think*. Morgan Kaufmann Publishers Inc. 1999.
- [9] Chi, E. H., & Card, S. K. (1999). Sensemaking of Evolving Web Sites Using Visualization Spreadsheets. In *Proc. of InfoVis'99*, pp. 18-26.
- [10] Holten, D. (2006). Hierarchical Edge Bundles: Visualization of Adjacency Relations in Hierarchical Data. *IEEE Transactions on Visualization and Computer Graphics*, 12, pp. 741-748.
- [11] Dou, W., Jeong, D.H., Stukes, F., Ribarsky, W., Lipford, H.R., & Chang, R. (2009). Recovering Reasoning Process from User Interactions. *IEEE CGA*, 29(3), pp. 52-61.
- [12] Fekete, J.-D., Wang, D., Dang, N., Aris, A. & Plaisant, C. (2003). Overlaying Graph Links on Treemaps. In *Proc. of InfoVis'03*, Poster Compendium, pp. 82–83.
- [13] Ghani, S., Elmqvist, N., & Ebert, D. S. (2012). MultiNode-Explorer: A Visual Analytics Framework for Generating Web-based Multimodal Graph Visualizations. . In *Proc. of EuroVA'12*, pp. 67-71.
- [14] Gotz, D. and Zhou, M. Characterizing Users' Visual Analytic Activity for Insight Provenance. In *Proc. of VAST'08*, pp. 123-130.
- [15] Gou, L. & Zhang, X. L. (2011). TreeNetViz: Revealing Patterns of Networks over Tree Structures. *IEEE TVCG/InfoVis'11*, 17(12), pp. 2449-2458.
- [16] Groth, D. P. & Streefkerk, K. (2006). Provenance and Annotation for Visual Exploration Systems. *IEEE TVCG*, 12(6), pp. 1500-1510.
- [17] Heer, J., & Boyd, D. (2005). Vizster: Visualizing Online Social Networks. In *Proc. of InfoVis'05*, pp. 33-40.
- [18] Heer, J., Card, S. K., & Landay, J. A. (2005). Prefuse: A Toolkit for Interactive Information Visualization. In *Proc. of ACM CHI'05*, pp. 421-430.
- [19] Heer, J., Mackinlay, J., Stolte, C. & Agrawala, M. (2008). Graphical Histories for Visualization: Supporting Analysis, communication, and evaluation. *IEEE TVCG*, 16(4), pp. 1189-1196.
- [20] Heer, J., Viégas, F.B. & Wattenberg, M. (2007). Voyagers and Voyeurs: Supporting Asynchronous Collaborative Information Visualization. In *Proc. ACM CHI'07*, pp.1029-1038.
- [21] Henry, N. & Fekete, J. D. (2006). MatrixExplorer: a Dual-Representation System to Explore Social Networks. *IEEE TVCG*, 12(5), pp. 677-684.
- [22] Henry, N., Fekete, J. D. & McGuffin, M. J. (2007). NodeTrix: a Hybrid Visualization of Social Networks. *IEEE TVCG*, 13(6), pp. 1302-1309.
- [23] Jankun-Kelly, T.J., Ma, K.L. & Gertz, M. (2007). A Model and Framework for Visualization Exploration. *IEEE TVCG*, 13(2), pp. 357-369.
- [24] Kadivar, N., Chen, V., Dunsmuir, D., Lee, E., Qian, C., Dill, J., et al. (2009). Capturing and Supporting the Analysis Process. In *Proc. of VAST '09*, pp. 131-138.
- [25] Kang, H., Plaisant, C., Lee, B., & Bederson, B. B. (2007). NetLens: Iterative Exploration of Content-Actor Network Data. *Information Visualization*, 6(1), pp. 18-31.
- [26] Klein, G., Moon, B., & Hoffman, R. R. (2006). Making Sense of Sensemaking 2: A Macrocognitive Model. *IEEE Intelligent Systems*, 21(5), pp. 88-92.
- [27] Kreuseler, M., Nocke, T. & Schumann, H. (2004). A History Mechanism for Visual Data Mining. In *Proc. of InfoVis'04*, pp. 49-56.
- [28] Larkin, J., & Simon, H. (1987). Why a Diagram is (Sometimes) Worth Ten Thousand Words. *Cognitive Science*, 11(1), pp. 65-100.
- [29] Liu, Z., Navathe, S. B., & Stasko, J. T. (2011). Network-based Visual Analysis of Tabular Data. In *Proc. of VAST'11*, pp. 39-48.
- [30] Martin, W. (2006). Visual Exploration of Multivariate Graphs. In *Proc. of ACM CHI '06*, pp. 811-819.
- [31] Meng, C., Yasue, M., Imamiya, A. & Mao, X. (1998). Visualizing Histories for Selective Undo and Redo. In *Proc. 3rd Asia Pacific Computer Human Interaction*, 1998. pp. 459-464.
- [32] Neumann, P., Schlechtweg, S. & M. S. T. (2005). ArcTrees: Visualizing Relations in Hierarchical Data. In *Proc. of EuroVis'05*, pp. 53–60.
- [33] Norman, D. (1993). Things That Make Us Smart: Defending Human Attributes in the Age of the Machine: Basic Books.
- [34] O'Madadhain, J., Fisher, D., Smyth, P., White, S., & Boey, Y. B. (2005). Analysis and Visualization of Network Data Using JUNG. *Journal of Statistical Software*, pp. 1-35.
- [35] Perer, A., & Shneiderman, B. (2008). Integrating Statistics and Visualization: Case Studies of Gaining Clarity during Exploratory Data Analysis. In *Proc. of ACM CHI'08*, pp. 265-274.
- [36] Perer, A., & Shneiderman, B. (2008). Systematic Yet Flexible Discovery: Guiding Domain Experts through Exploratory Data Analysis. In *Proc. of IUI '08*, pp. 109-118.
- [37] Pirolli, P., & Card, S. (2005). The Sensemaking Process and Leverage Points for Analyst Technology as Identified through Cognitive Task Analysis. *2005 International Conference on Intelligence Analysis*.
- [38] Pretorius, A. J., & Wijk, J. J. V. (2006). Visual Analysis of Multivariate State Transition Graphs. *IEEE TVCG*, 12(5), pp. 685-692.
- [39] Qu, Y., & Furnas, G. W. (2005). Sources of Structure in Sensemaking. In *Proc. of ACM CHI '05*, pp. 1989-1992.
- [40] Russell, D. M., Stefik, M. J., Pirolli, P., & Card, S. K. (1993). The Cost Structure of Sensemaking. In *Proc. of ACM INTERACT '93 and CHI '93*, pp. 269–276.
- [41] Shen, Z., Ma, K. L., & Eliassi-Rad, T. (2006). Visual Analysis of Large Heterogeneous Social Networks by Semantic and Structural Abstraction. *IEEE TVCG*, 12(6), pp. 1427-1439.
- [42] Shneiderman, B., & Aris, A. (2006). Network Visualization by Semantic Substrates. *IEEE TVCG*, 12(5), pp. 733-740.
- [43] Shrinivasan, Y. B. & van Wijk J. J. (2008). Support the Analytical Reasoning Process in Information Visualization. In *Proc. of ACM CHI'08*, pp. 1237-1246.
- [44] Stasko, J., Görg, C. & Liu, Z., (2008). Jigsaw: Supporting Investigative Analysis through Interactive Visualization. *Information Visualization*, 7(2), pp. 118-132.
- [45] van Ham, F., H.-J. Schulz & Dimicco, M. J. (2009). Honeycomb: Visual Analysis of Large Scale Social Networks. In *Proc. of INTERACT '09*, pp. 429-442.
- [46] Wright, W., Schroh, D., Proulx, P., Skaburskis, A. W., & Cort, B. (2006). The Sandbox for Analysis: Concepts and Evaluation. In *Proc. of ACM CHI'06*, pp. 801-810.
- [47] Zhang, J., & Norman, D. A. (1994). Representations in Distributed Cognitive Tasks. *Cognitive Science*, 18, pp. 87-122.
- [48] Zhang, X., Qu, Y., Giles, L. C., & Song, P. (2008). CiteSense: Supporting Sensemaking of Research Literature. In *Proc. of ACM CHI '08*, pp. 677-680.