TMNVis: Visual analysis of evolution in temporal multivariate network at multiple granularities

B. Lu\textsuperscript{a}, M. Zhu\textsuperscript{a,*}, Q. He\textsuperscript{a}, M. Li\textsuperscript{b}, R. Jia\textsuperscript{a}

\textsuperscript{a} Sichuan University, China
\textsuperscript{b} RMIT University, Australia

A R T I C L E   I N F O

Article history:
Received 16 September 2016
Revised 17 January 2017
Accepted 23 March 2017
Available online 12 April 2017

Keywords:
Temporal multivariate network
Network evolution
Visual analysis
Timeline-based method
Topological structure

A B S T R A C T

Temporal (Dynamic) multivariate networks consist of objects and relationships with a variety of attributes, and the networks change over time. Exploring such kind of networks in visualization is of great significance and full of challenges as its time-varying and multivariate nature. Most of the existing dynamic network visualization techniques focus on the topological structure evolution lacking of exploration on the multivariate data (multiple attributes) thoroughly, and do not cover comprehensive analyses on multiple granularities. In this paper, we propose TMNVis, an interactive visualization system to explore the evolution of temporal multivariate network. Firstly we list a series of tasks on three granularities: global level, subgroup level and individual level. Secondly three main views, which rely mainly on timeline-based method while animation subsidiary, are designed to resolve the analysis tasks. Thirdly we design a series of flexible interactions and develop a prototype system. At last we verify the effectiveness and usefulness of TMNVis using a real-world academic collaboration data.

1. Introduction

Networks are ubiquitous. Many real-world phenomena can actually be considered as networks, such as telecommunication, migration, academic cooperation. A network is defined as a series of nodes (i.e., vertices) connected by links (i.e., edges). In most cases, it also incorporates a variety of attributes associated with nodes and links. Such topological structure and attributes of network can change over time. Due to its time-varying and multivariate (multiple attributes) nature, above type of network is called temporal multivariate networks \cite{1}. It is critical to understand this type of network as the incorporated huge insights. For example, in a company’s mail system, nodes and links represent employees and their interactions respectively. The nodes have multiple attributes, e.g., name, age, department, position, and the links have some attributes, e.g., time, type (send or receive). Exploring the change of individuals’ activities can help managers catch the anomalies, which is critical for security concern.

Although a lot of visualization techniques have been used to explore evolution of network in multiple domains \cite{2–5}, current methods can not well meet the demand of exploring temporal multivariate network. Many researchers pay more attention to various aspects on network topology, such as detecting communities and analyzing their evolution \cite{6}. However, few of them focus on exploring multiple attributes in the process of network evolution. For example in academic cooperation, we not only want to know how does the number of researchers’ partners change along time (topology), but are also interested in whether the number of productive partners influences their development (attribute). We believe that more insights will be gained from exploration on topology and multiple attributes simultaneously. Additionally, different analysis granularities contain different information, such as overall structure in the entire network \cite{7}, formation and evolution of groups \cite{8} and evolution of node/link \cite{9}. However, few systems attempt to provide the comprehensive network analysis at multiple granularities, and most of existing systems only focus on one or two granularities, which makes users can not easily adjust their analysis granularity to gain more insights. Above all, it is of significance to provide a flexible and comprehensive system to help users explore the temporal multivariate network.

In this paper, we propose an interactive system called TMNVis to address the problems raised above. Based on the existing literature \cite{9–12}, we list a series of tasks upon three levels: 1) global level, 2) subgroup level and 3) individual level. At the global level, we use MDS \cite{13} and force-directed \cite{14} layout to analyze the evolution of attributes similarity and network topological structure respectively. Users are allowed to observe nodes of interest

\textsuperscript{*} Corresponding author.
E-mail addresses: thubinbo0326@163.com (B. Lu), zhumin@scu.edu.cn (M. Zhu), hjq163.com (Q. He), Mingzhao.Li@rmit.edu.au (M. Li), jiaruyouy0319@163.com (R. Jia).

http://dx.doi.org/10.1016/j.jvlc.2017.03.003
1045-926X/© 2017 Elsevier Ltd. All rights reserved.
in both layouts simultaneously. At the subgroup level, a glyph-based timeline visualization is designed to analyze evolution of subgroups and relationships between them. Last at the individual level, a timeline-based visualization method integrated with multiple subviews, is used to explore evolution of ego-network, which is formed with a specific node (called ego) and nodes connect to it (called alters). The above three views can be switched to each other easily with the assistance of a series of user interactions. Our main contributions are as follows:

1. We propose an interactive visualization system to explore the evolution (topology and attributes) of temporal multivariate network at three levels: global level, subgroup level and individual level.
2. We design a flexible and novel Subgroup View at the subgroup level and a timeline-based visualization integrated with multiple subviews at the individual level.
3. We conduct case studies with realistic academic collaboration network to evaluate the usefulness and effectiveness of TMNVis.

The paper is organized as follows. Section 2 will discuss the related work. Next, the analysis tasks will be described in Section 3. In Section 4 we will introduce our visual design and implementation. Then a case study on the dataset extracted from DBLP collaboration data will be given to verify the effectiveness and usefulness of our system in Section 5. We will give the discussion in Section 6. Finally, the conclusions and future work will be presented in Section 7.

2. Related work

In this section, we will describe related work from two aspects: visualization of dynamic network and visualization of multivariate network, which are two main aspects in our work.

2.1. Visualization of evolution in dynamic network

A lot of researchers have paid attention to the evolution of dynamic network using visualization. We will discuss them from two perspectives: analysis techniques and analysis granularity.

From the perspective of analysis techniques, existing work can be divided into two categories: animation-based (time-to-time mapping) and timeline-based (time-to-space mapping). Several papers [15,16] have given a comprehensive survey on them. Timeline-based methods provide a better overview of networks, and they are usually used with node-link diagrams to present network evolution intuitively and flexibly. We commonly place the network snapshots next to each other. To make the full use of space, some studies [17,18] position the nodes vertically on y-axis which is also used in this paper. However, timeline-based method is a kind of space-consuming technique. When it is leveraged to analyze dynamic networks with a little large scale or complicated structure, it has negative effect on the readability of network. Therefore, it is necessary to find a supplemental method, such as animation, to adapt to certain scenarios. Animation transition is firstly proposed by Eades and Huang [19] to explore changes of graph structure over time. In the process of animation, nodes and links dynamically appear, disappear or move to new position. This method helps us capture the changes of underlying topological structure more easily and save the space simultaneously. Bash et al. [11] used animation with multiple transition stages and temporal navigation to identify and track changes of network. We refer to their method in this paper to reduce users’ visual burden.

For the perspective of analysis granularity, most of the existing systems only focus on one or two types of granularities. For example, Vehlow et al. [20] developed a visualization approach to reveal typical life time phenomena of communities (subgroup). EmailTime [21] integrates both statistic and visualization to mainly analyze the activities of single email address (individual). Some studies [22,23] focus on the individual level, and the global level analysis is used as a simple sum of the individual level observations. There exists a few systems covering three analysis granularities mentioned before, such as TimeMatrix [10], GraphDiaries [11]. However, TimeMatrix is not suitable for some tasks as its limitation of layout and the matrix-based method in it is not intuitive enough compared with node-link diagram; while GraphDiaries focuses on network topology evolution lacking of exploration on multiple attributes.

2.2. Visualization of multivariate network

Multivariate networks always have a lot of attributes associated with nodes and links. A lot of methods have been proposed to explore the multiple attributes in network. One of the most general methods is visual encoding, namely, attributes are encoded as visual variables, such as color, size, shape [24–26]. Sometimes attributes will be used to compute attribute-based layout, such as in JauntyNets [27] and RSP [7], to provide more insights. Besides, multiple attributes also can be used to directly define a layout. For example, PivotGraphs [28] presents an view on the network by positioning nodes on x-axis and y-axis with categorical node attributes and the nodes with same attribute value will be aggregated. Similarly, Stef et al. [29] used two dimensions to directly define the position of nodes using a scatterplot and superpose the edges. The advantage of this method is that we can freely adjust the layout according to our analysis tasks.

However, most of the existing studies on exploring multiple attributes are mainly in static network, and there are less in dynamic network because of its dynamic characteristic. The researches on dynamic network mainly focus on topology and most of them only use the method of visual encoding to explore the multiple attributes. For example, Bach et al. [11] only used node size to represent the node degree in the process of exploration on network evolution. A few studies on dynamic network have relatively comprehensive exploration on multiple attributes. For example, Wu et al. [9] used a diversity of methods to explore these attributes, such as computing attributes-based layout, using rich visual variables. However, they mainly focused on ego-network, many tasks in this paper cannot be resolved well, especially at the subgroup level.

In summary, most of the existing systems do not support to analyze the temporal multivariate network comprehensively from diverse granularities and lack of exploration on multiple attributes. Compared with them, TMNVis provides a more comprehensive analysis on multiple granularities. Furthermore, a series of methods are provided to make full use of multiple attributes to gain more insights.

3. Task analysis

As shown in Fig. 2, we firstly extract network from the raw data. Then with user interactions, the extracted data will be processed in the data analysis module. Later it is transferred to different views in visual design module, and each view corresponds to specific tasks. The analysis tasks play an important role in the system as data analysis and visual design modules all depend on it. Therefore, this section we will introduce the analytical tasks module.

As talked above, existing systems lack of comprehensive analysis at multiple granularities. Yi et al. [10] believed that social network tasks should be supported at three levels: 1) temporal changes at the global level, 2) temporal changes at the subgroup level and 3) temporal associations among nodal and dyad level attributes. Ahn et al. [12] used three dimensions: entity, property
and temporal feature, to describe network evolution tasks. The first dimension, entity, consists of node-link, group and network, which follows the definitions of [10]. We use the same analysis granularities of Ahn, which are represented as individual, subgroup and global respectively in this paper. Although Ahn’s task taxonomy is very detailed, it is too complex and not consistent with our purpose of solving common tasks in different domains.

Referring to the task taxonomies on dynamic network [11,12] and other literatures [9,10,30], we list a series of tasks on temporal multivariate network. The tasks in different levels are independent and focus on different research objects. What’s more, the tasks don’t cover domain-specific tasks. For example, some network data has geographic information and the tasks may always highly relate to it. However, we are more likely to provide common tasks from topology and multivariate perspectives for different domains. Although the following tasks not cover all of the common tasks, but are representative.

**Global Level tasks.** At this level, we focus on evolution of overall network topological structure and nodes similarity (attributes-based). The tasks are listed as follows:

**Q1:** How does the overall topological structure change over time or what is the evolution pattern of entire network? Will the density of network become dense or the opposite?

**Q2:** How does the distribution of nodes change when using attributes-based layout? Are there coming to be a series of clusters? Whether there will be some outliers in the graph or not?

**Q3:** What is the evolution trend of the nodes with similar attributes? Whether they keep the same patterns or not?

**Subgroup Level tasks.** At this level, we focus on the evolution of groups (defined by attributes) and relationships between them. The tasks are listed as follows:

**Q4:** What is the evolution pattern of a specific subgroup? How does the size of subgroup change along time? How does the attribute distribution of individuals change over time?

**Q5:** How does the relationship change between subgroups? Whether the relationship is becoming stronger or not?

**Q6:** Are there any difference between subgroups? Such as the evolution trend, the nodes incorporated in the subgroup.

**Individual Level tasks.** At this level, we focus on evolution of individuals and relationships between them in an ego-network. The tasks are listed as follows:

**Q7:** How does the overall change of alters? Whether the number of alters is becoming larger or not?

**Q8:** How does the relationship strength change between a specific alter and the ego? Does it keep in a stable state?

**Q9:** What is the difference of evolution between alters? Whether the changes of attributes present the same trend or not?

**Q10:** Based on network connectivity, how does the cluster form with alters and separate over time?

4. Visual design and implementation

As shown in Fig. 1, the system consists of four components. In this section we will first give a brief description of design rationales. Then, we will describe the design and implementation of three panels: Evolution panel, Detail panel and Entity list. The main user interactions used in this paper will be listed at last.

4.1. Design rational

Based on task analysis, we compile the following design rationales:

- **Guide users’ exploration process:** To help users have a smooth exploration process from global to individual level, we would like to provide users with auxiliary views and flexible switch operation.

- **Allow exploring different attributes**: The exploration on attributes is critical in our system, especially at the subgroup level and individual level. We would like to enable users to operate attributes flexibly.
• Reduce visual clutter: A large amount of nodes and links can easily result in visual clutter, which makes users can not capture insights clearly. We would like to use different flexible interactions to resolve the problem as far as possible, e.g., filtering, sorting.

4.2. Evolution panel

4.2.1. Global view

According to the tasks listed at the global level, we use MDS layout, as shown in Fig. 1-b (top), to present distribution of nodes based on attributes similarity at each time step and then juxtapose the snapshots next to each other. Moreover, to perceive the topological structure evolution of the network intuitively, we use another layout, force-directed layout accompanied with animation as shown in Fig. 1-b (bottom).

MDS layout. Every entity is represented as a dot in the MDS. The distance between two dots represents the similarity of them, and the more similar the dots are, the smaller the distance is. We use a list of attributes of each entity $u$, including topology attributes and other multiple attributes, which can be used to describe the characteristics of $u$, to compute the layout of each time step. The topology attributes we used are as follows:

- Degree centrality: $\frac{n}{\sum_{i=1}^{N} c_i}$, where $n$ and $N$ denotes the u’s degree number and nodes number of entire network respectively.
- The average degree centrality of $u$’s neighbors: $\frac{\sum_{i=1}^{n} w_{e_i}}{\sum_{i=1}^{n} n_{u}^{e_i}}$, where $c_i$ denotes the degree centrality of the i-th neighbor.
- Average weight of the edges between $u$ and its neighbors: $\frac{\sum_{i=1}^{n} w_{e_i}}{\sum_{i=1}^{n} n_{u}^{e_i}}$, where $w_{e_i}$ denotes the weight of edge $e_i$.
- The density of edges between $u$’s neighbors or the clustering coefficient: $\frac{\sum_{i=1}^{n} w_{e_i}}{n(n-1)}$, where $|E_u|$ denotes the edge number between u’s neighbors.

Besides of topology attributes, we also use other attributes which are related to specific dataset, including numeric and categorical type. Now every entity can be described as a feature vector which consists of the attributes refered above. Then we use Canberra Distance [31] to compute the similarity between two entity $u$ and $v$:

$$\text{Similarity}(u, v) = \sum_{i=1}^{n} \frac{|F_u^i - F_v^i|}{(|F_u^i| + |F_v^i|)}$$

where $n$ denotes the number of features and $F$ denotes the feature vector. Compared with Euclidean distance or other distance measurement methods, this method helps us save the normalization step and computing time.

To fulfill the analysis tasks well at this level, we provide flexible selection to select nodes of interest. Through a brush, the entity list or search box, all the selected nodes will be highlighted at each time step. Moreover users can choose whether to show the lines that connect the same node across time steps through keyboard event ‘Shift+5’. Besides above, we allow users to use node color to present attribute that they are interested in.

Force-directed layout. Force-directed layout is usually used with node-link diagram to show graph structure. Every entity is represented as a node, and the node size is degree. In this paper we use FR algorithm [32] to position the nodes and links and use animation to track changes. To avoid confusing situation and rendering problem when users deal with large scale dataset, we allow users to filter out data according to their attributes. Thus they can focus on nodes of interest.

To help users gain insights of evolution, we use multiple stages to track the transition and highlight changes with different colors, as shown in Fig. 3(a-c). The transition stages are 1) Delete stage helping us capture the nodes or links that will be deleted soon with red highlight, 2) Position adjustment stage adjusting the remaining nodes/links with grey to their new location, 3) Add stage showing the new nodes/links with blue highlight. Furthermore, to keep the stable layout in the process of animation transition as far as possible, we utilize the stable state of previous time step as the initial state of the current time step.

Moreover, to reduce the memory burden, we use snapshots to record every time step of network, as shown in Fig. 1-b(bottom-right). We remain the state of the network at each time step, and users can easily forward or back to a special time step through the snapshots. Dragging the control bar of time-axis also has the same effect.

Similar to the selection in MDS, to track the changes of a specific node or group, we allow users to select nodes of interest using a brush as shown in Fig. 3-d. All the selected nodes will be highlighted with a red ring and remained in the process of animation transition (Fig. 3-e). What’s more, these nodes will also be highlighted in MDS simultaneously. This correlation presentation greatly help users gain a big picture of network evolution.

4.2.2. Subgroup view

Based on a timeline and flexible interactions, the Subgroup View is designed to analyze the evolution of subgroups. Users are allowed to divide the entire network into multiple subgroups based on nodes or links attributes through the Control panel. Each subgroup represented as a glyph (Fig. 4-a), consists of individuals belong to special attribute value or range. There will be no overlapping nodes in subgroups while using nodes attributes, and the edges between the subgroups denote the aggregated links between nodes in subgroups. Different from the former, when links attributes are used, there will be overlapping nodes in different subgroups and the edges indicate that there exists common nodes in the linked subgroups. Subgroups of different time steps are located horizontally along time, and different subgroups in each year will be located vertically. By default, all the horizontal lines are
positioned from top to bottom according to sum of subgroup size (amount of individuals) on the whole time.

Fig. 4-b presents the detail of glyph. Each rect (Fig. 4-c) presents a subgroup. The subgroup size (number of individuals) in each year is encoded as background color (Fig. 4-d) with a linear gradient and the greater the value is, the deeper the color is. The bar chart embedded in each rect (Fig. 4-c) indicates the distribution of the specific node attribute in a subgroup, and numeric and categorical type are both supported. For numeric attributes, users are allowed to divide the value domain into multiple fragments. The height of each bar indicates the number of nodes with corresponding attribute value in the subgroup. The proportion line, as shown in Fig. 4-e, indicates the proportion of new emerging nodes and continuous nodes. The continuous nodes in current time step also occur in the previous time step but the new emerging nodes do not. These two types of nodes are represented as green (new) and pink (continuous) respectively. The edge (Fig. 4-f) between subgroups indicates the relationship between them as talked before. To strengthen visual perception of users, dual mapping, color and thickness, is leveraged to illustrate the strength of relationship. The deeper the color is and thicker the edge is, the stronger the relationship is.

To gain more information of subgroups, users can flexibly toggle attribute that the bar chart shows, through double-clicking the embedded bar chart or the Control panel. For numeric attributes, users are allowed to divide the value domain into multiple fragments represented by bars with the same color. For the categorical attributes, fragments are encoded as different colors. To capture the change of a specific fragment of a subgroup over time, the corresponding fragments at different time steps will be highlighted.
when we focus on it. What's more, we increase the width of focused fragment and decrease the width of left fragments to avoid the situation that it is too thin to perceive the changes.

Moreover, to avoid visual clutter, we hide all edges by default. When we focus on a specific subgroup, only the edges and subgroups related to it will be shown. The edges can also be filtered by the widget in Control panel. In addition, we use an order method to solve overlapping of edges caused by limitation of space between adjacent subgroups on horizontal line. The steps of method are as follows: 1) Firstly compute the edge strength between focused subgroup and other subgroups on the whole time, and order them from small to big. 2) Secondly position the focused subgroup to the middle of the view. 3) Finally position the left subgroups on two sides of focused subgroup vertically, and the subgroup with stronger relationship will be located further.

To further capture the details, when we focus on the special subgroup, the Detail panel will show the top ten nodes that remain the longest time in the subgroup. It helps us select someone of interest to explore in the Individual View.

Besides the design above, we ever designed another alternative. As shown in Fig. 5, the glyph that presents a group, consists of a pie (Fig. 5-b) and a circle ring (Fig. 5-b). The circle ring is divided to indicate the proportion of new emerging nodes and continuous nodes with different colors and the pie is used to show the attribute distribution that user choose. In addition, the size of group is mapped to outer radius (Fig. 5-c). The visual design of edges between groups is the same with the design above. Compared to former one, though it is more easily to capture the proportion of attribute distribution, it makes a little visual confusion because of different size of glyphs, which is the main reason that we choose the former design.

4.2.3. Individual view

At the individual level, we use the four components shown above, to explore the individuals and relationships between them in an ego-network.

The bar chart (Fig. 6-a) depicts the overall change of the amount of alters over time as the height of bars indicates. The flow chart (Fig. 6-c), is designed to present the evolution of the attributes of alters. Users can easily select the attributes they are interested in through checkboxes. To navigate the process of exploration well, a spiral chart (Fig. 6-d) based on spiral-based layout is designed. The middle node indicates the ego and the distance between alter and ego indicates the strength of relationship on the whole time. Lastly the node size is encoded as the influence or the importance of alters.

The alter view (Fig. 6-b) is the main component in Individual View. Every small rect filled with color denotes an alter, and the color represents attributes. These rects are positioned vertically at each time step. Additionally, the bezier curve links the same alter in adjacent time steps to help users track the alter of interest. We allow users to change the color encoding and sorting strategy through flexible user interactions. As for color encoding, by default we use green and grey to indicate new emerging nodes and continuous nodes as defined in Subgroup View. Besides of this mapping scheme, the color can be used to demonstrate nodes attributes or links attributes using linear mapping. The larger the value is, the deeper the color is. For the layout of alters on y-axis, we provide three methods from different aspects: 1) by the intensity of cooperation with ego (links attribute), 2) by alters’ attributes (nodes attributes) and 3) by amount of individuals in cluster which is defined by network connectivity. The alters with same attribute value or in the same cluster will be bundled into a block as shown in Fig. 6-e, and different blocks are separated by space. To reduce the intersection of lines, the alters in each block will be ordered by a similar sorting method in [9]. The rules are as follows: 1) position the continuous nodes before the new emerging nodes and remain the order in the previous time simultaneously, 2) the new emerging alters that will still exit in the next time step will be ordered before. Besides of above operations, we provide a brush for users to select a group of nodes of interest, which is especially important when users analyze the cluster’s formulation and separate.

We also ever experimented with another alternative for alter view, as shown in Fig. 7. In this design, the same individual in different timesteps are located horizontally and are connected by lines (Fig. 7-c). The thickness of line indicates the value of relation strength change compared to last timestep and the color presents the rate of change as the legend shows. The individual is presented as a glyph by a circle (Fig. 7-b) and a ring (Fig. 7-a). The color of the circle can be encoded as attributes both for numeric and categorical types. For numeric type, the mapping rule is shown as the legend in Fig. 7 and for categorical type, we use different colors. The thickness of the ring also can be used to present numeric attribute. Meanwhile the users can take full use of the position of the individual on vertical direction according to attributes which is similar with the design above. Compared to the former design, it can show more attributes simultaneously and capture the evolution of individual more clearly as they are in the same horizontal line. However, it is limited by scalability and can not layout by network connectivity vertically.

4.3. Detail panel

Detail panel consists of three parts corresponding to views of three granularities to provide some auxiliary information for users. The first part presents the added or deleted nodes (Fig. 1-c) for Global View at each time step. Some important nodes in a special subgroup will be shown in the second part and then users can focus on the one of interest to further explore it in the Individual View. Last a spiral chart (Fig. 6-d) will be shown in the third part to guide the exploration, just as introduced in Individual View.

4.4. Entity list

As shown in Fig. 1-d, all the entities in network will be listed in the Entity list. It is an important component as we can focus on entities of interest and then they will be highlighted in other views. By default two attributes will be shown in it, and users are allowed to change the present attributes freely. The attribute values are encoded as bars with color. The larger the value is, the wider the bar is. Additionally all the entities can be sorted by attribute flexibly, which helps users capture the information of interest.
4.5. User interactions

Flexible interactions can help users gain more insights from data and navigate the exploring process. Now main interactions used in this paper are listed as follows:

**Data filtering and brushing** Many widgets are listed to help users filter out data according to attributes of nodes and links. We also allow them to select a group of nodes through a brush to track their changes.

**Data Aggregation** To fulfill tasks and reduce the crowded situation caused by limited space, we allow users to adjust the time granularity to aggregate the dataset.

**Temporal Navigation** In animation, we provide flexible navigation to forward or back to a specific time step through the bar on time axis or the snapshot records, which greatly reduces users’ short-term memory burden.

**Sorting** The method of sorting has been applied in multiple places in this paper. In the Individual View, users can choose different sorting methods to change the position of alters according to analysis task. In Entity list, users are allowed to sort entities by attribute of interest.

**Zoom in/out** As the limitation of space, the elements in views are usually too small to observe the detail when users deal with large dataset, especially in Subgroup View. We provide zoom in/out to solve this problem. Thus, users can get overview and details of network evolution flexibly.

**Transition and Highlight** Smooth transition and highlight are both widely used in this paper to help users track the changes more easily.

5. Case study

In this section we use the academic collaboration data from 1985 to 2013 which is extracted from ArnetMiner database [33] to verify the effectiveness and usefulness of TMNVis. We select the papers published in journals and conferences on visualization,
computer vision and computer graphics from above dataset. Finally 26,894 authors and 64,564 relationships are included.

After importing the dataset, we first explore it from Global View. Besides of the topology attributes reffered above, we compute the MDS layout with the following features:

1. p_num: publication number of the author at each year (dynamic, numerical).
2. t_pub: total number of publication on the whole timespan (static, numerical).
3. venue_type: the author publish paper on journal, conference or both of them at each year (dynamic, categorical).

Based on MDS layout (Fig. 8), we find that at the beginning only a few sparse clusters exit. With the time goes on, there comes to several stable clusters and simultaneously two relative large ones (Q2). We select a group of nodes in a small cluster using a brush as shown in Fig. 1-b, and we notice that majority nodes gradually transfer to the large clusters later, which is consistent with the whole evolution trend. To explore whether the similar authors have the same pattern, we select the top seven authors with high publication in the Entity list and find that six of them nearly have the same pattern (Q3) that they always in the same cluster. Only Shneiderman’s state is different from above six authors as he always occurs in other clusters over time, as the orange line (Fig. 8) shows.

Next we select the authors whose number of publication is larger than ten, to observe the change of their topological structure using animation. From Fig. 9, we can know that there are only some small groups appearing in early stage. With time goes on, more and more small groups connect to each other and finally become several relatively large clusters (Q1). We can infer that the authors from different groups have more cooperation than before.

From the Global View, we can find some interesting phenomena, such as most of the authors prefer to publish on conferences, which can be seen from MDS layout when we use different colors to present venue_type of node. However, we are interested in more insights. Therefore, we switch to the Subgroup View and the attribute of venue (link attribute, the name of journal or conference) is selected to divide the entire network into multiple subgroups, as shown in Fig. 10. Then the edge between two subgroups denotes the common authors. In this picture, we filter out the edges whose strength less than five, and select the node attribute of total number of publication for embedded bar chart. The value domain of the attribute is divided into three fragments: value ranging from zero to five, six to ten and above ten. We find that IEEE TVCG locates at the first line and then we focus on it as shown in Fig. 10-a. We notice that the IEEE TVCG first appears in 1995 and the number of its authors is increasing (Q4), especially from 2006 to 2013, which can be easily perceived from the darkness of background color. Besides, compared with early time, we notice that this journal has more continuous authors recently (Q4) when we zoom in to observe the proportion line as shown in Fig. 10-b. Meanwhile we find that more productive authors publish their work on IEEE TVCG than other venues (Q6), as embedded bar chart indicates. It is probably because it is a high quality journal approved by many researchers. Then we look at the edges related to IEEE TVCG and find that it keeps a strong relationship with EuroVis all the time. What’s more, the relationship has a tendency to increase (Q5). We can infer that authors who publish articles on IEEE TVCG prefer publishing on EuroVis than others.

Then we focus on the Journal of Visualization (JoV) to analyze different performance between IEEE TVCG and it (Q6). As shown in Fig. 11, JoV has less relationships with others, and the number of its authors is stable. What’s more, through the bar chart, we can find most of its authors have less publication records, which indicates these authors may be students.

To explore more details, we find that Bernd Hamann occur the most times in the subgroup of IEEE TVCG seen from the second part of the Detail panel. Additionally, he also has the most partners as shown in the Entity list. Then we switch to the Individual View to get more insights about him.

As shown in Fig. 12-a, we find that Bernd Hamann published paper in theses fields from 1990. In his active period (1996 to 2013), the amount of his partners presents a rapid growth especially from 1997 to 2007 and meanwhile he always had continuous partners every year (Q7), which can be seen from the grey rects. It may be caused by his related experience. Before 1995, Bernd Hamann was an assistant and associate professor at Mississippi State University, and then he was appointed by UC Davis which he stayed at for a long time. It makes him accumulate a number of stable partners. In addition, we can find there are always less than three connectivity-based clusters at each year except 2007, and most of clusters include many alters which indicates the partners of Bernd Hamann always have close ties with each other in that period. Then we select a cluster (Fig. 12-c) to observe how the cluster forms and how it develops. As shown in Fig. 12-b, we notice that a lot of partners included in this cluster
had ever occurred in previous clusters. In other words, it is not a new cluster but the one organized by previous clusters. Moreover some partners in this cluster had stayed together for a period of time through the observation of lines. However, the cluster disappears after 2004 (Q10), and only a few partners left and sparsely scattered in other clusters. Then we move to the spiral chart in Detail panel, and we can easily find Kenneth I. Joy kept cooperation with Bernd Hamann most (Q9). We select the top two alters and observe them by two sorting methods as shown in Fig. 13. We notice that Kenneth I. Joy kept a continuous relationship with Bernd Hamann from 1997 to 2005 and the number of publication keeps at a stable level, which can be seen from the alter's position in Fig. 13-a. Compared with Kenneth I. Joy, Hans Hagen did not keep a long continuous relationship with Bernd Hamann. Moreover, the number of his publication presents a decreasing trend as shown in Fig. 13-b. Then we order the alters by intensity of cooperation (Fig. 13-c and Fig. 13-d), and it shows that above two alters both kept a relative stable intensity of cooperation with ego (Q8). The exception is the cooperation between Kenneth I. Joy and ego increases sharply in 2000.

6. Discussion

Above case study demonstrates the effectiveness and usefulness of TMNVis to a large extent. In the process of exploration, we find it is enjoyable that we can select nodes optionally or directly position to a specific node of interest, and the context or related infor-
Fig. 11. Evolution of Journal of Visualization in the subgroup view.

Fig. 12. Bernd Hamann's Individual View. a) The alters are sorted by cluster. b) A cluster is selected and the alters in this cluster are also be highlighted in other time steps.
information will be highlighted immediately. Moreover, the changes can be easily gained through the rich visual encodings and attribute-based layout, especially for attributes in Subgroup View and Individual View.

However there are still many limitations in the exploration process. 1) First problem is the scalability. Due to the space limitation of timeline-based method, the number of time steps shown in Subgroup View and Individual View is better less than 30. Otherwise, the edges will be chaotic and the entire view will be crowded. Simultaneously, the number of nodes in force-directed layout is only suitable with hundreds of nodes. Although we provide a list of methods to weaken this effect, e.g., adjust time granularity to aggregate the data, we can not completely resolve the problem as some tasks can not be avoided. 2) Moreover, although we provide three main views to resolve the tasks of different granularities, we can not present three views simultaneously and users need to switch from one to another, which may result in losing the relevance between the views in their minds. 3) Furthermore, it is better to improve the MDS to construct a stable or consistent layout to strengthen the continued perception. 4) Finally, the linear mapping is not suitable to all attributes since many of them have different distributions. We need a better design for attribute mapping considering the distribution of each attribute.

7. Conclusion and future work

We have designed an interactive visualization system named TMNVis to explore the evolution of temporal multivariate network. A series of tasks are listed firstly, which cover three granularities: global level, subgroup level and individual level. We mainly used timeline-based method accompanied with animation to explore evolution of topological structure and multiple attributes. Meanwhile we provided rich and flexible interactions, including data filtering and brushing, data aggregation, temporal navigation, etc., to assist users in exploring the data. Finally, the academic collaboration data is used to verify the effectiveness and usefulness of our system.

In the future, we will use other user interactions and visual designs to display large scale dataset. Moreover, more distance measure methods, such as DDQC, KS, and attributes-based layouts, such as PCA, t-SNE, will be leveraged in system to strengthen analysis ability. Furthermore, we will validate the tasks proposed above with experts and add more compound tasks simultaneously. Finally, further to prove the effectiveness and usefulness of the system, more evaluations are needed.

Acknowledgment

We appreciate the suggestions from Dr. Shen. This work was supported in part by Sichuan Science and Technology Agency under Grant No. 2013GZ001, and the Technology Bureau of Chengdu under Grant No. 2014-HM01-00302-SF.

References


