A Task Taxonomy of Network Evolution Analysis

Jae-wook Ahn, Catherine Plaisant, and Ben Shneiderman

Abstract—Visualization is a useful tool for understanding the nature of networks. The recent growth of social media requires more powerful visualization techniques beyond static network diagrams. One of the most important challenges is the visualization of temporal network evolution. In order to provide strong temporal visualization methods, we need to understand what tasks users accomplish. This study provides a taxonomy of the temporal network visualization tasks. We identify (1) the entities, (2) the properties to be visualized, and (3) the hierarchy of temporal features, which were extracted by surveying existing temporal network visualization systems. By building and examining the task taxonomy, we report which tasks have been covered so far and suggest additions for designing the future visualizations. We also present example visualizations constructed using the task taxonomy for a social networking site in order to validate the quality of the taxonomy.

Index Terms—Network visualization, temporal evolution, task taxonomy, design space

1 Introduction

Network visualization is a crucial tool for understanding the nature of social networks. It can show the players of the networks and their relationships visually and can let the users explore and achieve important information for the tasks such as uncovering influential actors, finding helpful bridging people, and identifying destructive spammers. Existing social network analysis (SNA) software packages such as UCINET and Pajek support network visualization features. The recent growth of social media [5] requires more powerful social network analysis and visualization techniques beyond the conventional static network diagrams. One of the most important challenges is the visualization of temporal network evolution.

Time series visualization in general helps us to discover relations and patterns [3]; learn from the past to predict, plan, and build the future [1]. Therefore, various attempts have been made to provide effective tools for time series analysis. TimeSearcher [6, 15] provided an interactive pattern search in time series. Hochheiser and Shneiderman [18] introduced timeboxes which supported direct manipulation for specifying query constraints on time series data sets. Aris et al. [3] explored different strategies time intervals, specifically focusing on unevenly-spaced time series. Lifelines [33] and Lifeflow [35] provided methods to understand temporal categorical patterns. In social network analysis, the importance of temporal network analysis and longitudinal network models has been pointed out too [34]. However, this domain has been relatively less explored by social network researchers.

In order to provide strong temporal visualization methods, we need to learn which temporal tasks should be accomplished using the tools and which temporal aspects have been used so far. Yi [37] provided a classification of temporal visualization tasks and a list of measures for temporal visualizations. However, his taxonomy did not provide a full listing of possible temporal visualization tasks. Palla et al. [24] listed six possible types of community events but they did not provide any rationale or evidence to justify the classification. Fekete and Plaisant [12] defined general tasks for trees and Shneiderman [30], Amar [2] and Lee [20] presented comprehensive list of graph visualization tasks and relevant examples but they did not address the time dimension.

In this study, we propose a taxonomy of the temporal network visualization tasks. Within the taxonomy, we identify the entities and the properties to be visualized and showed the hierarchy of the temporal features extracted by surveying the existing temporal network visualization studies. By comparing the tasks and the visualization study examples, we can find out which tasks have been covered so far and suggest additions for designing future visualizations. At the same time, we apply the task taxonomy to a social networking site called Nation of Neighbors and provide network evolution analysis tasks on a design space.

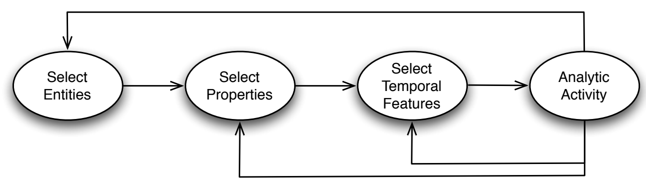

Fig. 1. Temporal network analysis process

The following section introduces three dimensions for defining the temporal network visualization. A list of tasks identified by combining the dimensions and their taxonomy is provided in Section 3. In Section 4, we propose the network evolution analysis tasks and show examples of the usage of the task taxonomy in a Nation of Neighbors social network analysis project. The last section concludes the paper and reveals our future plans.

2 Dimensions of Temporal Network Evolution Tasks

We identified the tasks, which are comprised of the combinations of the following triples: `[Entity-Property-Feature]`. The entities are the objects in which analysts are interested. For example, we can be interested in the growth (features) of node (entities) degrees (properties) while observing the age (properties) of each node. We found different network granularity of analysis could be adopted for selecting the entities (Section 2.1). When the entities are selected, the properties of the entities should be examined. The properties include the structural measures frequently used for network analysis and the domain attributes analyzed and compared over time (Section 2.2). Finally, we can select the temporal features specifically important for the temporal analysis (Section 2.3). In fact, the entities and the properties are the main elements of the conventional (static) network analysis too. However, we need to clearly identify the temporal features in order to understand the temporal nature of network evolution.

For the entire analysis task, users can iterate the selection of the triples according to the sub-tasks repeatedly. The iteration can be done for all the triples or only for a part of them as in Figure 1. During the iteration, users can combine independent tasks sequentially to form a larger compound task, too.

- *Jae-wook Ahn and Ben Shneiderman are with Human Computer Interaction Lab & Department of Computer Science, University of Maryland, College Park, MD, E-mail: {jahn,ben}@cs.umd.edu.*
- *Catherine Plaisant is with Human Computer Interaction Lab, University of Maryland, College Park, MD, E-mail: plaisant@cs.umd.edu.*

2.1 Entities of Analysis – Node, Group, and Network

The visual analysis of network evolution starts from the selection of the granularity or the level of analysis. By selecting different granularity, we can analyze different levels of temporal activity of networks. Yi [37] classified the tasks supporting the temporal social network visualization techniques into three levels: (1) Analysis of temporal changes at the global level, (2) Analysis of temporal changes at the subgroup level, and (3) Analysis of temporal associations among nodal and dyad level attributes. It is the classification about the scope of the entities – global network, group, or node – that we want to analyze and it conforms with our intuition about understanding the different network objects. Sometimes we are interested in observing individual player's activities while extending our scope of observation to a group of players or the network as a whole.

Therefore, we adopted these three levels as three entities that the users of the temporal social network visualization are interested in. It decides whether the user is interested in the temporal change of global network topology, the change of any specific attribute of subgroups, or node/links. The subgroups are defined as the intermediate entities between the entire network and individual nodes, such as triads, network motifs [22], and clusters. They can be sub-divided into two types: (1) **Structural Groups** and (2) **Domain Groups** according to the features to define them. The former is similar with the notion of *Subgroups* in [34] where the structural positions of the members decides the groupings. The latter is closer to *Social Groups* where the similarity of node attributes decides the groupings. We chose the terms Structural and Domain Groups in order to make clear the mechanism by which the groups are formulated. Likewise, the network level can be classified into two types: (1) **Connected Network** and (2) **Disconnected Components**. The former network is comprised of actors who are connected to each other while the latter includes disconnected separate sub-networks.

2.2 Structural Properties and Domain Attributes

Each entity – network, group, or node – can have a number of properties that the users are interested to compare over time. We broadly grouped them as (1) **Structural Properties** and (2) **Domain Attributes**. The Structural Properties reflect the topological nature of the entities. They include the general graph theory-based measures that are used for social network analysis. The latter defines any other kind of information that an network entity can have, regardless of the structural property.

Lee et al. [20] defined a graph visualization task taxonomy and classified the tasks as (1) Topology-based (adjacency, accessibility, common connection, connectivity, attribute), (2) Attribute-based (node and link attributes), (3) Browsing, and (4) Overview. Shneiderman and Aris [30] defined a collection of challenges as (1) Basic networks (2) Node/Link labels, (3) Directed networks, and (4) Node/Link attributes. They then identified eight basic tasks that can be covered by the Basic networks and incrementally added more tasks according to the increase of challenge level. The Structural Property in this study is equivalent to those of Topology-based of Lee's task taxonomy and the Basic network challenges of Shneiderman's taxonomy. It was defined for analyzing the temporal change of the properties that can show the topological or structural characteristics of the entities. There are a large number of structural properties and it is not the aim of this paper to provide a classification or a complete list of all the possible properties. The properties and the attributes used in the studies we surveyed are listed in Figure 2.

Domain Attributes are similar to the *attributes* in Lee's task taxonomy or *labels/attributes* of Shneiderman's taxonomy. Researchers frequently need to correlate the network structure (including its temporal change) with another dimensions and the Domain Attributes can work as hypothetical independent or dependent variables. Examples are conversation topic, geo-location, and demographic information of actors. The attributes used in the surveyed studies are listed in Figure 3.

2.3 Temporal Analysis Features

Temporal Analysis Features define how we can analyze the objects and the properties selected in the previous section. The granularity of analysis and the property are about deciding *what* to analyze. Selecting temporal analysis features is about *how* to observe, identify, or compare them over time. These features are the core of the temporal analysis whereas the former elements are valid with static networks too. For example, we can just observe a betweenness centrality (Structural Property) of a user (Node) in a conversation network and find out s/he is a person with higher prestige with in the network. However, if we are interested in temporal evolution, we need to observe the change of the property over time or compare the different values of the property of multiple time points. Here, the temporal changes (growth or contraction) that are the objectives of the comparing/observing actions are defined as the temporal analysis features.

We grouped the temporal analysis features into two broad categories according to the data type of the temporal events: (1) Individual events and (2) Aggregated events. The former is mostly about individual categorical events occurring in separate time points whereas the latter is dealing with the aggregated ordered set of individual time points.

2.3.1 Individual Event Features

1. **Single Occurrences** – Most fundamental temporal events occurring independently from others. For example, addition or deletion of nodes in a specific time point can be related to this feature.

2. **Birth** and **Death** – A special case of single occurrence that indicates a beginning or an ending of an event.

3. **Replacement** – Replacement can be simply defined as the sum of one deletion and another addition. The change of edge direction can be classified to this event. That is, deletion of edges in a specific direction plus edges to the opposite direction (or to bi-directional) .

2.3.2 Aggregated Event Features – Shape of Changes

Aggregated event features are related to time periods that take place for relatively longer durations than the discrete events. We identified five of the continuous changes that can illustrate the shape of the timeline. Gregory [15] proposed three categories in a similar context but we added more features for the network analysis.

1. **Growth** or **Contraction** – Shows whether an entity property increases or decreases over time. This can include the size of the entities as well as any specific property values or statistics. That is, the growth of a network is understood as the increase of the number of its nodes and edges.

2. **Convergence** or **Divergence** – Any measure or attribute can show changes over time in its initial stage but gradually becomes stable. That is, the amount of change converges to zero in a specific time point. Conversely, a stable state can become unstable and show increasing changes over time.

3. **Stability** – There is no or little change over time.

4. **Repetition** – The repetition of specific patterns over time. It can *Fluctuate* or show *Ritual* behaviors.

5. **Peak** or **Valley** – Whether an entity property increases or decreases suddenly.

2.3.3 Aggregated Event Features – Rate of Changes

The rate of changes was separated from the previous two categories because it is concerned about a different dimension. It is about measuring and comparing the amount of changes in a given specific time period. Moody [23] called this as *relational pace* and defined three different aspects. We took two of the classes that we considered more common.

1. **Speed** – Represents the static amount of change in a given period of time, with respect to levels.

2. **Acceleration** – Represents the change in speed.

3 Building the Task Taxonomy

By combining the three dimensions discussed in the previous section, we can now build the list of network evolution analysis tasks. The classification of the tasks is similar to that of the temporal features but we added two more classes: (1) temporal data processing tasks and (2) compound tasks. The data processing tasks are not directly linked to the temporal visualization *features* such as the growth or convergence over time but as important for analyzing the changes afterwards. Amar [2] included tasks such as Retrieve Value, Filter, and Compute Derived Value in his ten analytic tasks for graph visualizations. Our tasks play similar roles but they are more focused on the temporal changes. We did not repeat generic data processing tasks here. The compound tasks are mixtures of atomic tasks and built by combining those atomic tasks in order to present more complicated analytic processes.

Along with the task descriptions, we included examples (in the bullet lists) extracted from 15 real systems and studies. Table 1 summarizes the examples and their application domains. Figure 2 and 3 organizes the examples along with the temporal features and the entity properties/attributes respectively. The study/system keys were marked in the parentheses of the examples below.

3.1 Temporal Data Processing Tasks

1. **Decide the time scale of data** – This is the most basic task that almost every temporal network analysis should perform [4]. It decides whether the time interval of the visualization should be monthly, weekly, or daily-based, etc.

2. **Aggregate the raw data to the bigger scale** – According to the decided time scale, we need to aggregate raw data into bigger time scale. For example, aggregate hourly data to daily scale.

3. **Filter the raw data or sample continuous time to discrete points** – Retrieve a subset of temporal data from the entire set or sample discrete time points from the continuous data.

    - Divide the entire time series into two parts before and after a specific date (GeoTemporalNet, SocialDynamicsVis, MobiVis, Prajna).
    - Extract some interesting time points and compare them (SocialAction).
    - Retrieve sliding (overlapping) windows of temporal events (SoNIA).

4. **Decide the relativity and the alignment point of time** – Decide whether the time scale is relative or absolute. At the same time, decide the reference time point with which the remaining time points will be aligned.

3.2 Individual Temporal Event Tasks

1. **Single occurrences**

    (a) Examine a specific value of an entity of one or more discrete time point(s).

    (b) Compare the value of entities of multiple time points.

    (c) Compare the value of entities among entities.

    (d) Compare multiple time points using similarity measures.

    (e) Compare the events with domain attributes.

    - Observe the change of network structure on a specific date by overlaying the two cliques on a single node-link diagram (GeoTemporalNet).
    - Compare two sub-groups of callers (who were assumed to switched their cellphones) using structural equivalence measures (GeoTemporalNet, SocialDynamicsVis).
    - Compare multiple network diagram snapshots extracted from different time points by examining the change of edge weights of interest (SocialAction).
    - Identify a high concentration of cellphone calls on a specific date (SocialDynamicsVis).
    - Compare the temporal changes with geospatial changes (GeoTemporalNet, SocialDynamicsVis, MobiVis).

2. **Birth and Death**

    (a) Find if and when a specific entity appears and disappears.

    (b) Find an emergence of a new network structure such as an interaction pattern, or sub-groups.

    (c) Compare the events with domain attributes.

    - Browse and find when and how often does a specific type of forum participants appear (Durant).
    - Find when a group of callers disappear (by the call frequencies) and when another group of callers appear (SocialDynamicsVis).
    - Observe if an existing sub-group (dis)appears on a specific date (Prajna).
    - Observe the network structure change and find if there is any new emerging sub-groups (SoNIA-1, SoNIA-2).
    - Observe the network structure change and find if there appears a new communication pattern (SoNIA-3).
    - Identify the birth of the communication groups (iQuest, TeCFlow).

3. **Replacement**

    (a) Find the change of entity properties.

    (b) Compare the event with domain attributes.

    - Discover the switch of cellphone ID's occurred on the same day (Prajna).
    - Find the switch of edge directions according to change of the communication pattern (SoNIA-3).

3.3 Aggregated Temporal Event Tasks

1. **Growth and Contraction**

    (a) Observe the value of an entity measure increases or decreases.

    (b) Compare the growth or the contraction of an entity between time points.

    (c) Compare the growth or the contraction pattern among entities.

    (d) Compare the events with domain attributes.

    - Observe the growth of the overall network (Durant).
    - Observe the growth of the co-author groups (C-Group).
    - Observe the growth of the communication groups (iQuest, TeCFlow).
    - Observe if the transitivity of the global network grows or contracts (SoNIA-1, SoNIA-2).
    - Observe if the reciprocity of the global network grows or contracts (SoNIA-1, SoNIA-2).

- Observe the growth and contraction of cell phone call frequency. (SocialAction).

2. **Convergence and Divergence**

   (a) Observe the value of an entity measure and find if and when it converges to a specific point.
   (b) In case of convergence, find if there appears a new structure at the point.
   (c) Compare the convergence states.
   (d) Compare the events with domain attributes.

   - Find if the transitivity converges and stabilizes after growing to a specific time point (SoNIA-1, SoNIA-2).
   - Find the convergence point of the transitivity and observe if the resultant network emerges (SoNIA-1, SoNIA-2).
   - Compare the result of the emerging network with the social balance theory. That is, whether the process of making friends is achieved through already close friends (SoNIA-1).
   - Find if there is any difference between the global network and its sub-groups in terms of the convergence metric (SoNIA-2).

3. **Stability**

   (a) Find if a changing value of an entity stabilizes.
   (b) Identify when the stabilization happens.
   (c) Compare the stability states.
   (d) Compare the events with domain attributes.

   - Observe if the collaboration pattern is growing or stabilized and compare them by international region (TimeMatrix).

4. **Repetition**

   (a) Find out if a pattern of an entity value change repeats.
   (b) Identify the repeating pattern.
   (c) Compare the repetition patterns.
   (d) Compare the events with domain attributes.

   - Observe the value of the network reciprocity measure fluctuates (SoNIA-2).
   - Observe the repeated communications between a teacher and his students (SoNIA-3).
   - Compare the two different communication patterns of two classrooms, one of which is more obedient and the other is not (SoNIA-3).

5. **Peaks** or **Valleys**

   (a) Find out if there is any peaks or valleys of an entity value change over time.
   (b) Identify the shape of the peaks/valleys. Do they change sharply or slowly?
   (c) Compare the peak/valley patterns.
   (d) Compare the events with domain attributes.

   - Identify a sudden peak within a time range and observe their duration is short or long. Compare them with structural properties or domain attributes such as topic (Shamma).
   - Find any number of collaboration of the players peaked. If any, when was it (TimeMatrix)?

3.4 Rate of Changes

1. **Fast** or **Slow**

   (a) Identify how much changes an entity had during a given time period.
   (b) Compare the difference of changes of multiple entities. Find out which one is faster or slower.
   (c) Compare the events with domain attributes.

   - Compare the speed of growth of nodes by different attribute types (Durant).

2. **Accelerating** or **Decelerating**

   (a) Identify whether a change is getting faster or slower.
   (b) Compare the acceleration or deceleration.
   (c) Compare the events with domain attributes.

3.5 Compound Tasks

There are special tasks that are the combinations of the single atomic tasks introduced so far. We can define almost endless number of compound tasks. Here, we show an example found from the studies we surveyed.

- (1) Identify two types of groups from the entire network +
  (2) Identify the starting time point of the collaboration network +
  (3) Identify the starting time point of the knowledge sharing network + (4) Compare the starting points and identify the sequence + (5) Compare the discovered sequence with the hypothesized sequence by domain knowledge (TimeMatrix).

3.6 Lessons Learned

The goals of building the network evolution task taxonomy are: (1) to learn about the common strategies of existing techniques and (2) to find out the probable future strategies that were not discovered by them yet. We can summarize the lessons learned from the task taxonomy that we have built so far as follows.

1. **Domain attributes prevail** – Almost all examples incorporated relevant domain attributes (Figure 3). This is rather a natural observation because studies usually include special domain attributes in their hypotheses and attempt to prove them by comparing the network evolution and the domain attribute values.

2. **Features less explored** – By mapping the study examples and the temporal network visualization features (Figure 2), we could identify the empty spots on the map (where no real system or study was placed) and could catch the clues for future additions. The most noticeable empty space is the *Rate of Changes* in the Aggregated Time Event Features column. According to our knowledge, Durant's was the only study that explicitly mentioned about the rate of changes (speed) in their real data analysis. For non-network time series visualizations, it is not a new topic (e.g. [29]) and the value of this feature for network visualization was already noted by [23]. However, real visualization-based temporal network analysis systems have not much supported this feature yet.

3. **Individual versus aggregated temporal events** – Almost all examples we examined used the individual temporal features as they are the most basic elements that should be analyzed. The aggregated temporal trend features were relatively less explored, except the rather simpler ones such as growth and contraction.

4. **Multiple granularity of analysis** – A lot of examples covered more than one entity. However, they were mostly in the node level analysis and accompanied the network level analysis as a simple sum of the node-level observations. Few studies attempted to provide users with means to control the granularity of visualization of analysis that can span the node/link, group, and the global network level.

Table 1. Temporal network visualization examples

Key	Authors	System/Study Name	Application Domain
Durant	Durant [10]	-	discussion board
GeoTemporalNet	Ye [36]	GeoTemporalNet	VAST08 cellphone network mini-challenge [16] [1]
SocialDynamicsVis	Farrugia [11]	SocialDynamicsVis	VAST08 cellphone network mini-challenge
MobiVis	Correa [9]	MobiVis	VAST08 cellphone network mini-challenge
Prajna	Swing [31]	Prajna	VAST08 cellphone network mini-challenge
SocialAction	Perer [25]	SocialAction	VAST08 cellphone network mini-challenge
SoNIA-1	Moody [23]	SoNIA	social network – social balance
SoNIA-2	Moody [23]	SoNIA	social network – Newcombs fraternity
SoNIA-3	Moody [23]	SoNIA	social network – education
iQuest	Gloor [14]	iQuest	communication archive (e-mail, phone records, blogs, etc)
TeCFlow	Gloor [13]	TecFlow	email archive
G-Group	Kang [19]	C-Group	citation network
Zhang	Zhang [38]	-	invitation network
Shamma	Shamma [27, 28]	-	microblog communication
Powell	Powell [26]	-	affiliation network of life science institutions

[1] Out of 23 VAST'08 mini-challenge participants, we sorted out these five teams as examples. Even though they had common elements to the solution, we listed them here independently due to the sub-task level differences in detail.

Fig. 2. Network Evolution Examples (Organized by the Temporal Analysis Features)

● Node/link, ⊗ Group, ❄ Network

System/Study keys	Temporal Analysis Features									
	Individual Time Event Features			Aggregated Time Event Features						
				Shape of Changes					Rate of Changes	
	Single Occurrences	Birth/Death	Replacement	Growth Contraction	Convergence Divergence	Stability	Repetition	Peak/Valley	Speed	Accelerate
Durant	●	●		❄					●	
GeoTemporalNet	● ⊗		● ⊗							
SocialDynamicVis	● ⊗	●	⊗							
MobiVis	● ⊗		⊗							
Prajna	❄	●	●							
SocialAction	●		●	❄						
SoNIA-1	●	⊗		❄	❄					
SoNIA-2	●	⊗		❄	❄					
SoNIA-3	●	❄			❄		❄			
iQuest		⊗		⊗						
TecFlow		⊗		⊗						
Shamma								❄		
Zhang				● ⊗						
TimeMatrix	⊗			❄		❄		●		
C-Group		⊗		⊗						
Powell	⊗ ❄	●		❄						

Fig. 3. Network Evolution Examples (Organized by the Entity Properties and Attributes)

System/Study keys	Entity Properties			
	Structural Properties		Domain Attributes	
Durant	●	Edge direction	●	Node type: provider,consumer,facilitator
GeoTemporalNet, SocialDynamicVis, MobiVis, Prajna, SocialAction	● ⊗ ❄	Degree, Edge direction/weight, Network layout	●	Call frequency, Geo-location, Cellphone ID
SoNIA	● ❄	Edge direction/weight, Transitivity, Reciprocity	❄	Class type: obedient, rebellious
iQuest	⊗	Betweenness-centrality	⊗	Event/knowledge creation, Origins of new idea
TecFlow	⊗	Betweenness centrality, Core/periphery structure	⊗	Innovation networks
Shamma	●	Centrality	●	Topic
Zhang	●	Degree, Hub	●	Organizational position, Acceptance rate
TimeMatrix	●	Density, Centrality	⊗ ❄	Inter-organizational collaboration
C-Group	●	Edge direction of the focal-pairs, Group membership	⊗	Author group
Powell	⊗ ❄	Cohesion, Homophily		

4 APPLYING THE TASK TAXONOMY – NATION OF NEIGHBORS EXAMPLE

This section shows examples of temporal network visualization for a social community site, called Nation of Neighbors (NoN). This illustrates how the taxonomy helps us plan for a variety of tasks we may not have thought of otherwise. At the same time, by comparing the guidelines for analyzing the NoN service and the task taxonomy we can check whether there is any missing task in the taxonomy.

NoN (`http://www.nationofneighbors.com`) is a web-based community network that enables neighbors to share local crime, suspicious activity, and other community concerns. It began in Jefferson County, WV, where it achieved a great success as "Watch Jefferson County." The NoN team has expanded their efforts across the U.S. in many communities. We are collaborating with them to help community managers explore and analyze the social dynamics embedded in their social networks.

We analyzed the community forum conversation network data provided by the NoN team.[2] As of February 2011, it contains conversations of 249 users in 137 communities, from November 2005 to February 2011 for around 5 years. There are 1,595 posts and replies within the dataset. We defined the following guidelines to best support the management of NoN.

1. NoN is a community-based network service. The network managers are interested in which communities are successful and what makes the differences. At the same time, they need to monitor the behavior of the entire network and should be able to find any anomalous behaviors of individual people. Therefore, multiple level of analysis comprising sub-groups, the global networks, and individual users is important (***Multiple granularity of analysis***).

2. In order to analyze the temporal evolution of the communities, the means to measure the various aspects of the networks are important. They include the structural metrics that can quantify the community activities (post, reply, crime report, page views and login/out) and the measure of success of the service.

   The users of the visualization (e.g. network managers) need effective methods to compare and analyze these metrics over time and understand the temporal changes. The comparative analysis should be accomplished visually in order to track the trends of a large amount of data easily (***Measures of activities and visual analysis methods***).

3. Along with the structural information, topic analysis is crucial to follow what is really happening inside the community, given the nature of the data (textual conversation in forums). This topic information needs to be visualized in a way that the users can understand the correlation between the topic and the structure change of the networks (***Domain attributes***).

4.1 Design Space

We can convert these guidelines to more specific tasks by referring to the task taxonomy. In order to make the process systematic, we adopted an idea of the design space [7, 8, 21]. Mackinlay [21] separated the data, graphical design, and the final rendered images, so that he could formally define the underlying requirements of the problem and decide the best presentation method (even automatically). In order to construct the taxonomy of input devices (e.g. mice, keyboards, or menus), Card and Mackinlay mapped various input devices into a design space that located the devices according to their core attributes. They located each input device according to their physical property (delta force, force, movement, or position) and to whether it has either a linear or a rotary dimension. For example, a menu is a device with a linear dimension (vertical or Y axis) and it has a position property. A mouse is a device with a linear dimension (vertical and horizontal or X and Y axis) with movement property.

[2] The user identities were anonymized to protect their privacy.

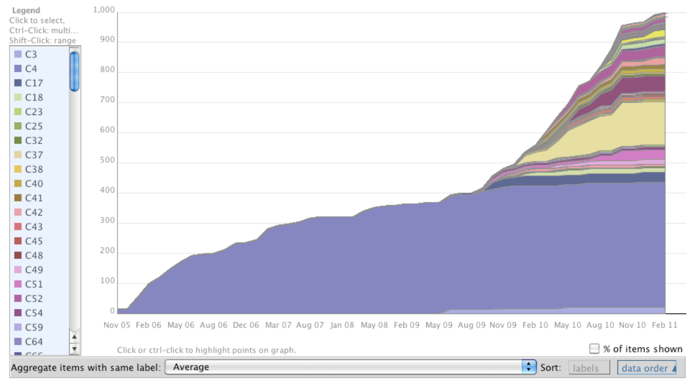

Fig. 5. Showing the growth of the entire NoN service network (all stacks) and individual communities (one per each stack)

The space has high expressiveness [21] in that it could map every input device into the space without making any additional incorrect mis-mapping. They could explain the individual nature of the input devices, find out the relationships among multiple devices, and even could suggest what future input devices should get equipped, by examining the empty spots – where no existing device was on the map.

Figure 4 is the design space of the temporal network visualization tasks for NoN. We show two dimensions on the map (1) Entities by the Granularity of Analysis (X axis) and (2) Temporal Analysis Features (Y axis). We did not show the last dimension (Structural Properties and Entity Attributes) in order to avoid the complexity that could be caused by integrating all three dimensions into a single 2-dimensional map.

In the following sections, we show the example visualizations that were built based on the design space for the NoN tasks. We selected the best example tasks from the design space that could satisfy the guidelines for NoN and confirm that the taxonomy could work for real social network visualization problems.

4.2 Visualization of Network Statistics in Different Granularity Levels

Figure 5 shows a stacked timeline visualization (generated using ManyEyes of IBM [32]). It illustrates two of the three entities from the taxonomy (Section 2) – sub-groups and the entire network. Each stack within the visualization represents a community and the sum of all stacks is equal to the growth of the global network. The blue stack (C4) that occupies the largest area is assumed to be the Jefferson County, which was the most successful one. From around November 2009, the entire network shows an exponential growth but the growth owes more to the other smaller communities than this major community, which shows just a logarithmic growth after November 2009. This example shows the importance of the multiple granularity of analysis and suggests that NoN can benefit from it.

4.3 Visual Analysis of Various Structural Properties

Figure 6 shows a prototype system called TempoVis. It was designed to show the temporal evolution of the networks using a node-link diagram combined with a "time-slider." Using the time-slider, users can navigate through time and see the snapshot of the network at any time point. In order to highlight the temporal changes, it emphasizes the nodes/links added at the *current* time point (month, in this example) in red colors and paints the past conversations in low intensity colors degrading by their ages. This function was derived from the idea of individual temporal events (Section 2.3.1 and Section 3.2) of the task taxonomy in order to contrast the information that cannot be found easily from the aggregated visualization as Figure 5.

The timeline graph on top of the time slider can show the shape of changes (Section 2.3.2 and Section 3.3) of the various network struc-

Fig. 4. Design Space of Nation of Neighbors Task Taxonomy

<table>
<tr><th colspan="3"></th><th colspan="3">Granularity of Analysis</th></tr>
<tr><th colspan="3"></th><th>Node/Link</th><th>Group</th><th>Network</th></tr>
<tr><td rowspan="29">Temporal Features</td><td rowspan="6">Individual Temporal Features</td><td rowspan="4">Single Occurrences</td><td colspan="3">Examine the number of community activities (post, reply, report, invitation, page view, login) at a time point</td></tr>
<tr><td colspan="3">Examine structural metrics (degree, density, centrality) at a time point</td></tr>
<tr><td colspan="3">Compare the number of activities and the structural metrics between time t1 and t2</td></tr>
<tr><td colspan="2">Compare the activities and structural metrics among users and communities</td><td></td></tr>
<tr><td>Birth/Death</td><td colspan="2">Find when a user/community activity appears/disappears</td><td></td></tr>
<tr><td>Replace</td><td>Find if and when the edge direction (replies) of users changes</td><td></td><td></td></tr>
<tr><td rowspan="18">Shape of Changes</td><td rowspan="3">Growth Contraction</td><td colspan="3">Observe the growth/contraction of the activities</td></tr>
<tr><td colspan="3">Observe the growth/contraction of the structure metrics</td></tr>
<tr><td colspan="2">Compare the activity growth among the users and the communities</td><td></td></tr>
<tr><td rowspan="4">Convergence Divergence</td><td colspan="3">Observe if a structure metric converges at a specific time point</td></tr>
<tr><td></td><td colspan="2">Find if a new structure emerges from the convergence</td></tr>
<tr><td></td><td colspan="2">Compare the convergence states between time points</td></tr>
<tr><td></td><td>Compare the convergence states among communities</td><td></td></tr>
<tr><td rowspan="3">Stability</td><td colspan="3">Find if the activities or the structural metrics are stable</td></tr>
<tr><td colspan="2">Compare the stability states among users and communities</td><td></td></tr>
<tr><td colspan="2">Compare the stability states between time points</td><td></td></tr>
<tr><td rowspan="4">Repetition</td><td colspan="3">Find if the activities or the structural metrics change pattern repeat</td></tr>
<tr><td colspan="3">Identify the pattern of the repetition</td></tr>
<tr><td colspan="2">Compare the repetition between time points</td><td></td></tr>
<tr><td colspan="2">Compare the repetition among users and communities</td><td></td></tr>
<tr><td rowspan="4">Peak/Valley</td><td colspan="3">Find if/when the activities or the structural metrics show a peak or a valley</td></tr>
<tr><td colspan="3">Identify the shape of the peaks/valleys</td></tr>
<tr><td colspan="3">Compare the position and the shape of the peaks/valleys between time points</td></tr>
<tr><td colspan="2">Compare the position and the shape of the peaks/valleys among users and communities</td><td></td></tr>
<tr><td rowspan="5">Rate of Changes</td><td rowspan="2">Speed</td><td colspan="3">Identify how much changes of the activities or the structural metrics occur at a give time</td></tr>
<tr><td colspan="3">Compare the speed across entities. Which one is faster or slower?</td></tr>
<tr><td rowspan="3">Accelerate</td><td colspan="3">Identify whether a change of the activities and the structural metrics is getting faster or slower</td></tr>
<tr><td colspan="3">Compare the acceleration between time points</td></tr>
<tr><td colspan="3">Compare the acceleration among users and communitiees</td></tr>
</table>

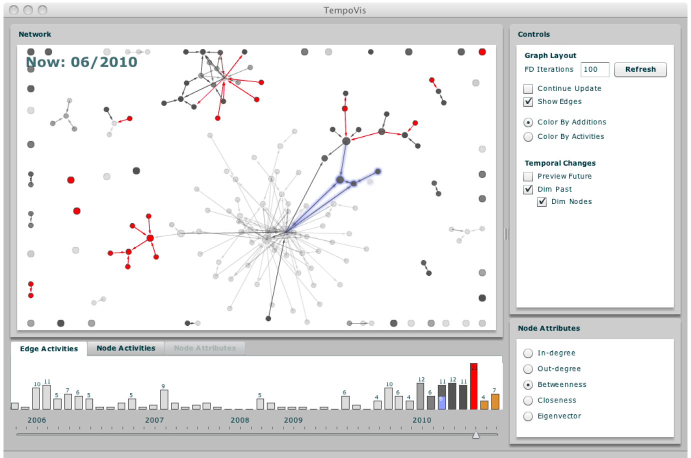

Fig. 6. TempoVis visualization. Shows dynamically updated node-link diagrams (above) according to the "time-slider." A time line graph (below) is linked to the slider and the visualization.

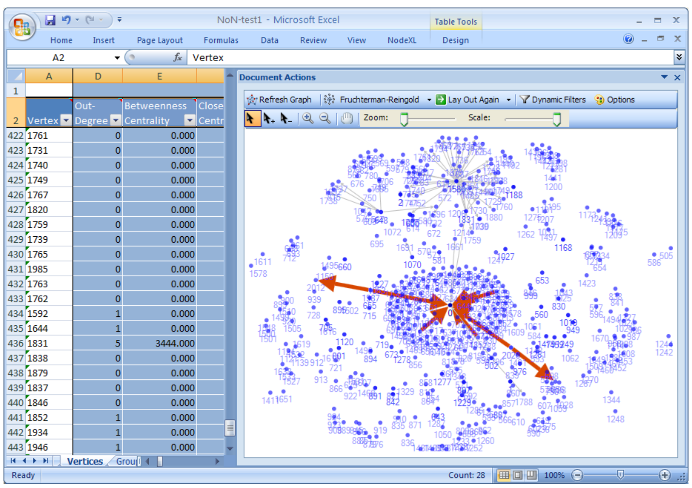

Fig. 7. NodeXL shows the temporal addition of nodes/links visually.

ture metrics defined in the taxonomy (Section 2.2) including in/out-degree, betweenness centrality, closeness centrality, and eigenvector centrality. Users can manipulate the time slider, observe the trend of the network property values changing, and compare them with the network visualization.

Figure 7 shows another example called NodeXL [17]. NodeXL is a Microsoft Excel add-on that can easily draw network visualizations from the data store in Excel worksheets. It supports various network layout algorithms, network structure metrics, clusterings, time-based filtering, and so forth. Using the basic functions provided by NodeXL, we could generate a visualization in Figure 7, which shows the temporal changes on a certain month in thicker arrows. These two examples show that the task taxonomy can cover the second requirement of NoN, by supporting the longitudinal comparisons of structural network properties using diverse interactive visualization methods.

4.4 Domain Attributes and Newly Discovered Temporal Features

The topics can be classified as a domain attribute and it is defined in the task taxonomy (Section 2.2). They can be analyzed from the textual conversation data extracted from the NoN online forums. We found the forum contents encompassed numerous topics, from criminal activities to everyday events around the neighborhoods. A domain expert is constructing a list of possible topics and related keywords by examining the contents and we are going to process the topic information automatically with the help of the topic list. Even though the visualization of the topics is still underway, we can find similar approaches such as [14].

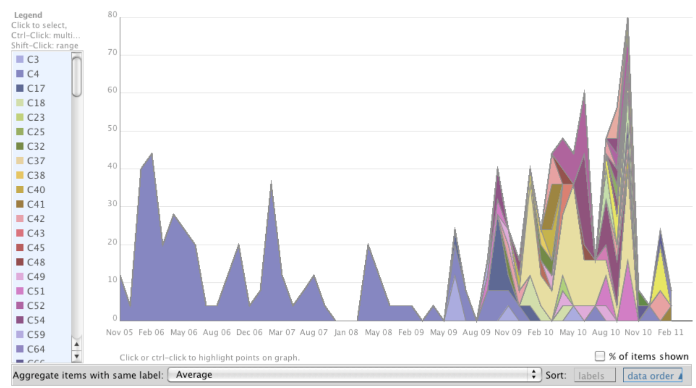

Fig. 8. The speed of changes of the entire NoN service network and individual communities

So far, we have examined if the task taxonomy could meet the requirements provided by the NoN team. We found the three requirements were successfully covered by the features listed in the taxonomy. At the same time, we found some unexpected temporal features that could be applied to the NoN problem. For example, Figure 8 shows the conversation growth speed per community and for the entire network. This is a new feature discovered by examining the task taxonomy developed independently from the NoN requirements. It shows the benefit of the task taxonomy that can discover new unknown features.

5 Conclusions and Future Work

This paper proposed a task taxonomy for network evolution analysis. We suggested to divide the definition of the tasks using three dimensions and identified the elements of each dimension. By combining these dimensions, we were able to formulate a task taxonomy based on previous work and visualization system examples. We could learn the features utilized so far and discover new aspects for the future development from the taxonomy. The task taxonomy provided us with several lessons: (1) the importance of domain attributes, (2) features less explored, (3) higher propensity to the simpler individual temporal features, and (4) the lack of means to incorporate different granularity of analysis.

For the future work, we are planning to incorporate more diverse domain attributes in addition to the topics, such as geo-spatial location of the communities in the NoN service. We believe that we could better understand the nature of the social networks by matching the various network features and the domain knowledge. With regard to the development of temporal network visualization systems, we will integrate the newly discovered visualization features to the existing NodeXL and TempoVis tools. It will be one of our future challenges to efficiently combine the variety of features discovered in this study and provide well-integrated user interfaces.

Acknowledgments

We appreciate the support of Microsoft External Research for the development of the free, open source NodeXL (`http://www.codeplex.com/nodexl`) and for National Science Foundation grant IIS0968521, Supporting a Nation of Neighbors with Community Analysis Visualization Environments, plus National Science Foundation-Computing Research Association Computing Innovation

Fellow, Postdoctoral Research Grant for Jae-wook Ahn. We appreciate the assistance of Art Hanson and the NodeXL team.

References

[1] W. Aigner, S. Miksch, W. Müler, H. Schumann, and C. Tominski. Visualizing time-oriented data–a systematic view. *Computers & Graphics*, 31(3):401 – 409, 2007.

[2] R. Amar, J. Eagan, and J. Stasko. Low-level components of analytic activity in information visualization. In *Proceedings of the 2005 IEEE Symposium on Information Visualization*, pages 15–21, Washington, DC, USA, 2005. IEEE Computer Society.

[3] A. Aris, B. Shneiderman, C. Plaisant, G. Shmueli, and W. Jank. Representing unevenly-spaced time series data for visualization and interactive exploration. *Human-Computer Interaction-INTERACT 2005*, pages 835–846, 2005.

[4] S. Bender-deMoll and D. McFarland. The art and science of dynamic network visualization. *Journal of Social Structure*, 7(2), 2006.

[5] d. boyd and N. B. Ellison. Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication*, 13(1):210–230, 2008.

[6] P. Buono, A. Aris, C. Plaisant, A. Khella, and B. Shneiderman. Interactive pattern search in time series. In *Proc. SPIE*, volume 5669, pages 175–186, 2005.

[7] S. K. Card, J. D. Mackinlay, and G. G. Robertson. The design space of input devices. In *Proceedings of the SIGCHI conference on Human factors in computing systems: Empowering people*, CHI '90, pages 117–124, New York, NY, USA, 1990. ACM.

[8] S. K. Card, J. D. Mackinlay, and G. G. Robertson. A morphological analysis of the design space of input devices. *ACM Trans. Inf. Syst.*, 9:99–122, April 1991.

[9] C. Correa, T. Crnovrsanin, C. Muelder, Z. Shen, R. Armstrong, J. Shearer, and K.-L. Ma. Cell phone mini challenge award: Intuitive social network graphs visual analytics of cell phone data using mobivis and ontovis. In *Visual Analytics Science and Technology, 2008. VAST '08. IEEE Symposium on*, pages 211 –212, 2008.

[10] K. Durant, A. McCray, and C. Safran. Modeling the temporal evolution of an online cancer forum. In *1st ACM International Health Informatics Symposium (IHI 2010)*, pages 356–365, November 2010.

[11] M. Farrugia and A. Quigley. Cell phone mini challenge: Node-link animation award animating multivariate dynamic social networks. In *Visual Analytics Science and Technology, 2008. VAST '08. IEEE Symposium on*, pages 215 –216, 2008.

[12] J.-D. Fekete and C. Plaisant. Infovis 2003 contest – general tasks applicable to most trees. `http://www.cs.umd.edu/hcil/iv03contest/generaltasks.html`, 2003.

[13] P. Gloor, R. Laubacher, Y. Zhao, and S. Dynes. Temporal visualization and analysis of social networks. In *NAACSOS Conference, June*, pages 27–29, 2004.

[14] P. Gloor and Y. Zhao. Analyzing actors and their discussion topics by semantic social network analysis. In *Information Visualization, 2006. IV 2006. Tenth International Conference on*, pages 130–135. IEEE, 2006.

[15] M. Gregory and B. Shneiderman. Shape Identification in Temporal Data Sets. Master's thesis, University of Maryland, 2009.

[16] G. Grinstein, C. Plaisant, S. Laskowski, T. O'connell, J. Scholtz, and M. Whiting. VAST 2008 Challenge: Introducing mini-challenges. In *Visual Analytics Science and Technology, 2008. VAST'08. IEEE Symposium on*, pages 195–196. IEEE, 2008.

[17] D. L. Hansen, B. Shneiderman, and M. A. Smith. *Analyzing Social Media Networks with NodeXL: Insights from a Connected World.* Morgan Kaufmann, 2010.

[18] H. Hochheiser and B. Shneiderman. Dynamic query tools for time series data sets: timebox widgets for interactive exploration. *Information Visualization*, 3(1):1–18, 2004.

[19] H. Kang, L. Getoor, and L. Singh. Visual analysis of dynamic group membership in temporal social networks. *ACM SIGKDD Explorations Newsletter – Special issue on visual analytics*, 9(2):13–21, December 2007.

[20] B. Lee, C. Plaisant, C. S. Parr, J.-D. Fekete, and N. Henry. Task taxonomy for graph visualization. In *Proceedings of the 2006 AVI workshop on BEyond time and errors: novel evaluation methods for information visualization*, BELIV '06, pages 1–5, New York, NY, USA, 2006. ACM.

[21] J. Mackinlay. Automating the design of graphical presentations of relational information. *ACM Trans. Graph.*, 5:110–141, April 1986.

[22] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon. Network motifs: simple building blocks of complex networks. *Science*, 298(5594):824, 2002.

[23] J. Moody, D. McFarland, and S. Bender-deMoll. Dynamic network visualization. *American Journal of Sociology*, 110(4):1206–41, 2005.

[24] G. Palla, A. Barabási, and T. Vicsek. Quantifying social group evolution. *Nature*, 446(7136):664–667, 2007.

[25] A. Perer and B. Shneiderman. Balancing systematic and flexible exploration of social networks. *IEEE Trans. Vis. Comput. Graph*, 12(5):693–700, 2006.

[26] W. Powell, D. White, K. Koput, and J. Owen-Smith. Network Dynamics and Field Evolution: The Growth of Interorganizational Collaboration in the Life Sciences. *American Journal of Sociology*, 110(4):1132–1205, 2005.

[27] D. Shamma, L. Kennedy, and E. Churchill. In the limelight over time: Temporalities of network centrality. In *Proceedings of the 29th international conference on Human factors in computing systems (CSCW 2011)*, Hangzhou, China, March 19-23 2011. ACM.

[28] D. Shamma, L. Kennedy, and E. Churchill. Peaks and Persistence: Modeling the Shape of Microblog Conversations. In *Proceedings of the 29th international conference on Human factors in computing systems (CSCW 2011)*, Hangzhou, China, March 19-23 2011. ACM.

[29] G. Shmueli, W. Jank, A. Aris, C. Plaisant, and B. Shneiderman. Exploring auction databases through interactive visualization. *Decision Support Systems*, 42(3):1521–1538, 2006.

[30] B. Shneiderman and A. Aris. Network visualization by semantic substrates. *IEEE Transactions on Visualization and Computer Graphics*, 12(5):733–740, 2006.

[31] E. Swing. Award: Efficient toolkit integration solving the cell phone calls challenge with the prajna project. In *Visual Analytics Science and Technology, 2008. VAST '08. IEEE Symposium on*, pages 221 –222, 2008.

[32] F. B. Viegas, M. Wattenberg, F. van Ham, J. Kriss, and M. McKeon. Manyeyes: a site for visualization at internet scale. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1121–1128, 2007.

[33] T. D. Wang, C. Plaisant, A. J. Quinn, R. Stanchak, S. Murphy, and B. Shneiderman. Aligning temporal data by sentinel events: discovering patterns in electronic health records. In *Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, CHI '08, pages 457–466, New York, NY, USA, 2008. ACM.

[34] S. Wasserman and K. Faust. *Social network analysis: Methods and applications*. Cambridge university press, 1995.

[35] K. Wongsuphasawat, J. Gomez, C. Plaisant, T. Wang, B. Shneiderman, and M. Taieb-Maimon. LifeFlow: Visualizing an Overview of Event Sequences. In *Proceeding of the twenty-ninth annual SIGCHI conference on Human factors in computing systems*. ACM, 2011.

[36] Q. Ye, T. Zhu, D. Hu, B. Wu, N. Du, and B. Wang. Cell phone mini challenge award: Social network accuracy – exploring temporal communication in mobile call graphs. In *Visual Analytics Science and Technology, 2008. VAST '08. IEEE Symposium on*, pages 207 –208, 2008.

[37] J. S. Yi, N. Elmqvist, and S. Lee. TimeMatrix: Analyzing Temporal Social Networks Using Interactive Matrix-Based Visualizations. *International Journal of Human-Computer Interaction*, 26(11):1031–1051, 2010.

[38] J. Zhang, Y. Qu, and D. Hansen. Modeling user acceptance of internal microblogging at work. In *CHI 2010 Workshop on Microblogging*, 2010.