

Visual Analysis of Dynamic Group Membership in Temporal Social Networks

Hyunmo Kang

University of Maryland
College Park, MD 20742
kang@cs.umd.edu

Lise Getoor

University of Maryland
College Park, MD 20742
getoor@cs.umd.edu

Lisa Singh

Georgetown University
Washington, DC 20057
singh@cs.georgetown.edu

ABSTRACT

C-Group is a tool for analyzing dynamic group membership in temporal social networks over time. Unlike most network visualization tools, which show the group structure within an entire network, or the group membership for a single actor, C-Group allows users to focus their analysis on a pair of individuals. While C-Group allows for viewing the addition and deletion of nodes (actors) and edges (relationships) over time, its major contribution is its focus on changing group memberships over time. By doing so, users can investigate the context of temporal group memberships for the pair. C-Group provides users with a flexible interface for defining (and redefining) groups interactively, and supports two novel visual representations of the evolving group memberships. This flexibility gives users alternate views that are appropriate for different network sizes and provides users with different insights into the grouping behavior. We demonstrate the utility of the tool on a scientific publication network.

Keywords: dynamic group detection, visual data mining, group visualization.

1. INTRODUCTION

There is a large and growing interest in tools that allow users to navigate, understand, manipulate and analyze the huge influx of complex, heterogeneous data with which today's data analyst is faced. The best of these tools combine powerful visualizations with useful analytics that are suited to the task at hand. Often by focusing on a specific analytic task, the visualization can be designed to highlight relevant information and support appropriate task-specific navigation and manipulation primitives. In this paper, we present C-Group, a visual analytic tool for pairwise analysis of dynamic group membership over time in two-mode affiliation networks. Given a pair of actors, referred to as the *focal pair*, C-Group enables an analyst to explore questions related to group membership of the focal pair, e.g. when are they in similar/different groups or how does the structure of the shared groups change over time?

One novelty of C-Group is its focus on the relationships and group memberships in which a pair of actors participate. This is in contrast to the majority of social network visualization tools which either show the whole network (but allow filtering and selection so that users can, through a series of operations, reach a desired sub-graph), or show an ego-centric view of the network, centered around a single individual. Our choice to focus on an actor pair was motivated by our earlier work on D-Dupe [5], a visual analytic tool for entity resolution in social network data. D-Dupe highlights pairs of actors which are very similar, and hence might be duplicates. It also shows a sub-network that highlights the neighborhood of the pair in a meaningful way for duplication detection.

While the pairwise view of C-Group was inspired by D-Dupe, because the analytic task that C-Group is designed to support is more complicated than entity resolution, significant changes and new developments were required in the design of the tool. One of the more straightforward changes was in the support for the choice of the focal pair; we allow users flexibility in how they search for the focal pair, either by direct search, similarity search, or a combination of the two. A second more fundamental design development was in the definition of groups; while this is often straightforward for networks describing a single entity type, in the context of an affiliation network, especially a dynamic affiliation network, there is a great deal more flexibility in how groups may be defined. Section 3 details our group semantics, but at a high level, a group represents a context specific collection of associates. This context may be based on node features, edge features, or network structural properties. Finally, C-Group is designed for longitudinal analysis of temporal social networks. All the time composition semantics for defining evolutionary periods are new contributions.

In terms of visualization, the major contribution of C-Group is its group context window, which shows the evolution of group membership for the focal pair over time. Most tools which support the visualization of changes in social networks over time are designed for the task of understanding the entire network, and are focused on the additions and deletions of nodes and edges in a single-mode social network. C-Group, on the other hand, is concerned with the evolution in relationships and group membership over time for the focal pair. This leads to two novel visualizations, one which highlights the change in the shared actors with which the focal pair associates, and another which highlights the group movements. In both cases, C-Group provides an appropriate stable layout, and uses animation in a way that truly brings attention to relevant portions of the changing network.

The paper is organized as follows. The next section presents related work. Section 3 describes the social network model used by C-Group. In Section 4, we introduce C-Group and give an overview of its features. Section 5 describes the focal pair selection, the construction of dynamic groups, and the visualization of the changing groups. We then present two case analyses of the tool, followed by conclusions and future directions.

2. RELATED WORK

Visual analysis of social networks is an integral component of the field of social network analysis and is in many ways fundamental to its very definition [6]. There are several excellent surveys describing social network visualization within the social sciences literature [7] and the information visualization literature [10]; in addition there many useful social network software packages that have limited visualization, but sophisticated statistical analysis [2][15] and others that focus more on interactive visualization of the complete network [1] [3] [9][12][13].

At the heart of most social network visualization work is understanding the groups that actors belong to and the role the actors play in these groups [7]. The majority of visualization tools accomplish this by showing single-mode networks with edges present between nodes for display of affiliations, and/ or color coding actors with different affiliations. Another approach is to optimize the layout by clustering like nodes or nodes that are in close proximity to each other [2]. Tools that consider heterogeneous networks are emerging [14]. However, the visual paradigms used for grouping remain largely the same as for single-mode models. In all these cases, the focus is on the entire network or on a single node's network, its ego-network. While analysis of pairs of nodes may be possible after a number of steps, the visualization is not optimized for visual analysis of pairs even though they are an important class within the context of social networks. For pair analysis, other tools require the user to be savvy enough to manipulate the visualization to clearly understand the group structure as it relates to the pair of interest.

There has also been a growing interest in visualizing dynamic social networks [11][4]. There are many interesting and complex issues involved both in terms of graph layout and the semantics of the dynamic networks. Again, the majority of that work has focused on single mode networks, such as friendship networks [8], and most of the work is interested in understanding the entire network, rather than a specific actor, or pair of actors.

While our work builds on all of these areas, for example our alternate visualization of the dynamic context graphs are related to the static flip books and dynamic movies of Moody et al. [11][4], our visualizations are geared toward understanding the evolution and changing composition of context specific dynamic groups. It is the targeted pairwise analytic task, the robust group semantics, and the novel visual paradigm for dynamic changes to group structure that distinguish our work from other visual analytic tools.

3. DATA MODEL

3.1 Temporal Social Network Representation

C-Group is used to visualize changing group membership in two-mode social networks, where there are two types of entities, Actors and Events, and there is a Participant relationship which links actors to events. The actors are often people, but they may be other social animals, webpages, etc. The events can, literally, be events such as an academic conference or a performance, and the participant relationship captures the attendance at the event, or, more abstractly, a co-occurrence of a set of actors, such as a publication, where the participant relationship is the coauthor relationship between an author and a paper. These are referred to as affiliation networks in the social network research community[16], and are often viewed as bipartite graphs, with actors as one set of nodes, events as the other set of nodes, and participation links between actors and events. Affiliation networks capture a wide variety of interesting application domains, including communication data (e.g. email and mobile phone calls) among people, organizational data describing people's roles on teams or in companies, and epidemiological data, describing people and the specific disease strain with which they are infected. Because our focus here is on temporal social networks, each event has some notion of 'time' associated with it. We support three different time compositions for longitudinal analysis:

- Time period: an established interval of time, e.g.) a year
- Time point + duration: a starting time point and a time period, e.g.) a start time + five minutes
- Multiple time points: a starting point and an ending point, e.g.) an observation start time and an observation end time

Each of the entities and relationships in the network may have additional associated attributes. It is straight-forward to describe an affiliation network using the following relational schema over actor relation \mathcal{A} , event relation \mathcal{E} , and participation relation \mathcal{P} .

- $\mathcal{A}(Id_a, A_1, A_2, \dots, A_m)$
- $\mathcal{E}(Id_e, E_1, E_2, \dots, E_n, E_{time})$
- $\mathcal{P}(Id_a, Id_e, P_1, P_2, \dots, P_p)$

Here, Id_a and Id_e represent unique identifiers for an actor and an event respectively. Id_a and Id_e together represent a unique relationship between an actor and an event in the participation relation. Notice that the event relation \mathcal{E} has a time attribute(s), E_{time} that will be used as the time granularity for the temporal analysis.

3.2 Grouping Semantics

As previously mentioned, one of the main contributions of this paper is the grouping semantics. Groups can be defined based on collections of actors or collections of events. Which approach is best will depend on the data set and the analytic task at hand; this in turn will influence the appropriate visualization.

3.2.1 Relational construct of groups

In C-Group, a group is always defined, extensionally, as a collection of actors. The composition of this collection of actors or group is defined based on shared attribute values of an actor, an event, or a participation relation. We now formally define three grouping constructs. For clarity, we will use a running example of a publication data set in which the actors are authors, the events are publications, and the participation relation is paper author:

- $Author(Id_a, First, Last, Affiliation)$
- $Paper(Id_e, Title, Year)$
- $Paper.Author(Id_a, Id_e, Role)$

Actor attribute grouping: For each specific value x of actor attribute A_i , we define the group $G_{A.A_i=x}$, where

$$G_{A.A_i=x} = \{a \mid a.A_i = x, a \in A\}$$

Here a group is defined based on a shared actor attribute that is not multi-valued. If two actors have the same actor attribute value, they are considered to be in the same group. Categorical attributes are most effective here. It is interesting to note that since these attributes are single-valued, they serve as a *partitioning* attribute for actors connected to the focal pair and, therefore, will be non-overlapping. For our example data set, we can define a group based on the affiliation of an author. In this case, our analysis would involve tracking the dynamics of shared affiliates for the focal pair.

Event attribute grouping: For each specific value y of event attribute E_j , we define the group $G_{E.E_j=y}$, where

$$G_{E.E_j=y} = \{a \mid \exists e p(a, e) \wedge e.E_j = y, a \in A, e \in E, p \in P\}$$

Here a group is defined based on a shared event attribute. Actors that participate together in an event and have the same event attribute value are considered members of a specific event attribute group. This type of grouping may have multiple values

for an actor. Therefore, we can view these as multi-valued attributes with potentially overlapping groups. For our example data set, we can define a group based on the topic of a paper. In this case, our analysis would involve tracking the dynamics of the topics as it relates the focal pair to other actors in the network.

Participation attribute grouping: For each specific value z of participation attribute P_k , we define the group $G_{P_k=z}$, where

$$G_{P_k=z} = \{a \mid \exists p(a, e) \wedge p.P_k = z, a \in A, e \in E, p \in P\}$$

Here a group is defined based on a shared participation attribute. Actors that participate together in an event and have the same participation attribute value are considered members of a specific participation attribute group. This type of grouping may also have multiple values for an actor, leading to potentially overlapping groups. For our example data set, we can define a group based on the role of participation, where role may represent whether an actor is a primary author or a secondary author. In this case, our

analysis would involve tracking the dynamics of the authors' role as it relates the focal pair to other actors in the network.

Hybrid attribute grouping: A group can be defined based on any combination of the above grouping constructs. In other words, a group can be defined using multiple basic grouping constructs connected via boolean AND and OR operators.

3.2.2 Dynamic group

The above provides a static definition of groups; because we are interested in the evolution of the groups, a temporal component is added to each grouping relation above. Recall that we support three different time compositions for temporal analysis. Therefore, we augment each of our grouping structures to exist at different times t : $G_{A.A_i=x}(t)$, $G_{E.E_j=y}(t)$, and $G_{P.P_k=z}(t)$. It is this time varying analysis of different grouping constructs that C-Group has been optimized to support.

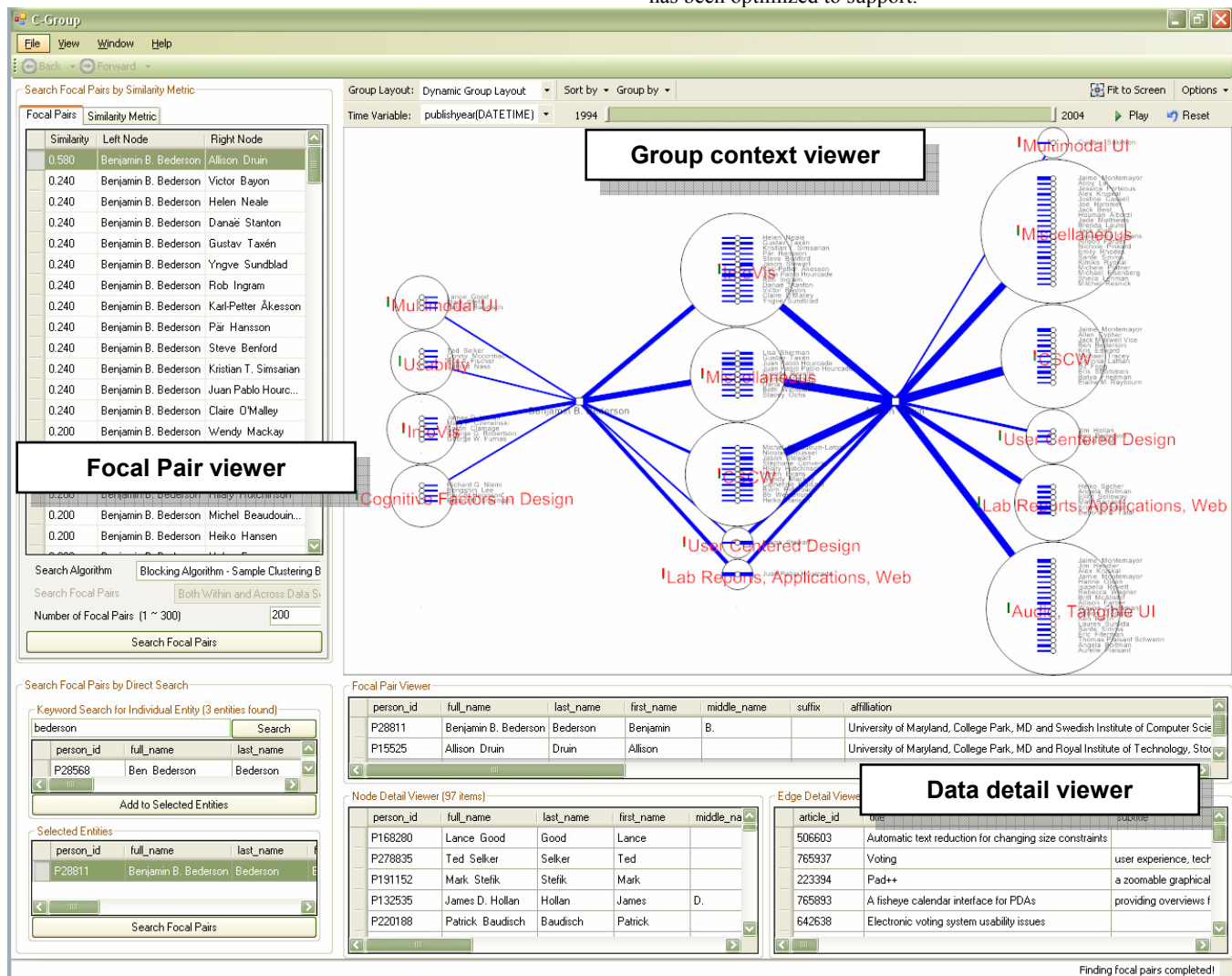


Figure 1. C-GROUP consists of three coordinated windows: the focal pair viewer on the left, the group context viewer in the upper right corner, and the data detail viewer in the lower right corner. The focal pair viewer allows users to search for focal pairs based on the users' interest. The group context viewer visualizes the group membership relation of the selected focus author pair using two different view modes: a fixed entity layout view and dynamic group layout view. The data detail viewer shows all the attribute values of the authors and papers displayed in the group context viewer.

4. OVERVIEW OF C-GROUP

In this section, we will begin our introduction of C-Group using a simple example with the publication relation described in the previous section. The dataset itself consists of a subset of the ACM Digital Library and contains 4,073 papers from the ACM-CHI conference from 1982 to 2004 authored by 6,358 people. Authors are connected with 12,727 co-authorship relation edges.

Figure 1 shows the overall C-GROUP interface, which is composed of three coordinated windows: the focal pair viewer, the group context viewer, and the data detail viewer. The focal pair viewer (on the left) allows users to navigate the affiliation network and select focal pairs of interest. Users can select a pair in the focal pair viewer, and then the group context viewer (on the upper right corner) provides two visualizations of the group membership context between the focal actors: the fixed entity layout and dynamic group layout. Figure 1 shows a dynamic group layout which represents the co-authorship relation between the two focal authors, "Ben Bederson" and "Allison Druin". A group of authors is represented as a circular node whose size is proportional to the number of authors it contains. In this example, all the shared coauthors of Ben and Allison are placed in-between them, which are grouped by the topics of the papers that were coauthored by Ben and Allison. On the other hand, non-shared coauthor groups are placed on the sides. For example, if an author is in the InfoVis group in-between Ben and Allison, it means that this author has a set of InfoVis papers coauthored by Ben and another set of InfoVis papers (sets can be overlapping) coauthored by Allison. Each edge between a focal pair author and a group represents all the papers co-authored by the people in the group and the focal pair author. The thickness of an edge represents the number of co-authored papers. With this simple, stable and meaningful layout, users can quickly get an overview of the shared and non-shared collaborators of the focal pair and their group membership.

As shown in Figure 1, each group node can display its individual nodes and their edges to support group analysis in more detail. The authors displayed in the viewer can be grouped either by author's attributes or by paper's attributes using the *Group By* menu in the tool bar. Groups are not necessarily mutually exclusive, which means, an individual actor node can be contained in multiple groups. For example, if an author has both InfoVis and CSCW papers coauthored with the focal pair authors, this author is a member of both InfoVis and CSCW groups. In addition to the dynamic grouping of authors, the changes of authors' group membership over time can be explored using the time range slider located in the toolbar. The construction of groups as well as the visualization of dynamic group membership evolution are described in more detail in the following section.

Finally, the data detail viewer (in the lower right corner) shows all the attribute values of authors and papers displayed in the group context viewer. The data detail viewer is tightly-coupled with the group context viewer so that users can easily choose nodes and edges in the network layout and observe the corresponding attribute values of any author or paper.

5. USER INTERFACE DESIGN

In this section, we discuss our design principles and explore the challenges in designing an interface for pairwise analysis of dynamic group membership by describing the main features of C-Group following the steps of the analysis process.

5.1 Focal Pair Selection

Recall that the visualization is optimized for analyzing the group membership dynamics of a *pair* of actors. We refer to this pair as the *focal pair* and refer to the two actors involved as actor 1 and actor 2. The focal pair is selected using the focal pair selection panel, which is shown on the left side of Figure 1. Users have a variety of options for choosing the focal pair, and the choice will depend both on the analytic task at hand, and how familiar users are with the data contained in their database.

Direct Search for the Focal Pair: Users may have very targeted analytic goals, in which they are interested in understanding the relationship among actors whose identities are known a priori. In this case, as shown in Figure 2, users simply search for the actors by keywords using the *Keyword Search* group box (located at the top of Figure 2) and then select and add the found entities to the *Selected Entities* list (on the bottom) by clicking the "Add to Selected Entities" button. By pressing the "Search Focal Pairs" button, users can see all the combinations of the selected entities as focal pairs in the focal pair viewer.

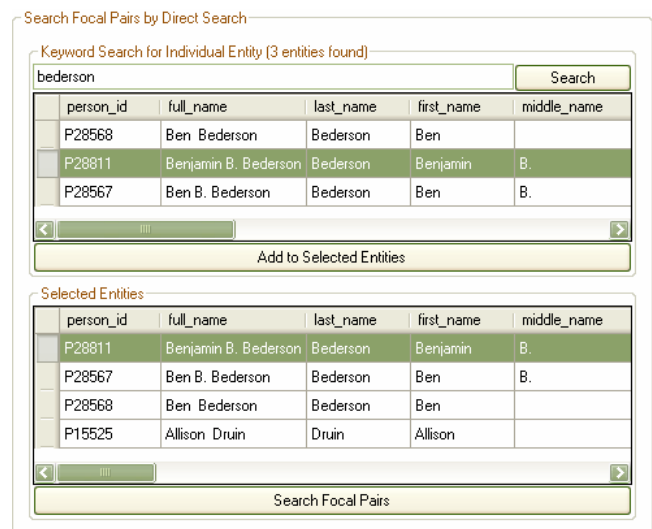


Figure 2. Direct search for focal pairs using keywords

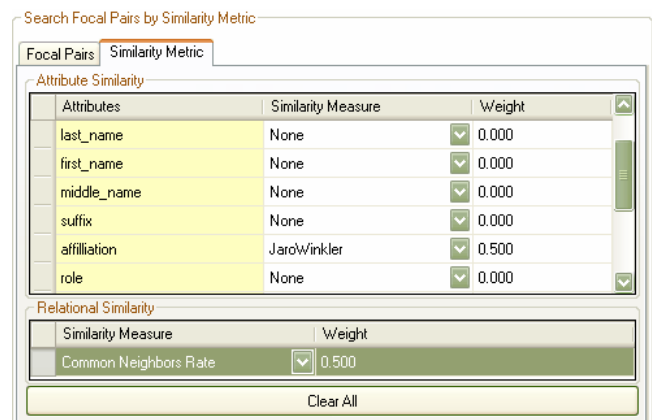


Figure 3. Search for focal pairs using a similarity metric

Similarity Search for a Focal Pair: Users with less targeted analytic goals, who are perhaps less familiar with the data, or are exploring to better understand the data, can use similarity search to find a focal pair. There is a great deal of flexibility in how

similarity is calculated. The decision can be based on similar attributes values, or, it can be based on the relational similarity between the actors. Figure 3 shows the similarity metric tab added to the focal pair viewer that enables users to define a similarity metric using both attribute and relational similarity. In the attribute similarity group-box, users can see all the entity attributes available in the imported data set in the first column. The similarity measure column allows users to choose a similarity measure function for each attribute. The choice of similarity measure function depends upon the application. C-Group is designed to enable users to explore multiple similarity function combinations and orderings easily and flexibly. In the last column, users can assign either positive or negative weights to the corresponding similarity functions. In Figure 3, JaroWinkler string match function is applied to the “affiliation” attribute and the Common Neighbors relational similarity is selected to search for the focal pairs that have similar affiliations and share many coauthors. Once the user has defined a similarity criterion, the database is searched, and pairs of similar actors are presented. The user can then select a pair of interest to explore further.

Combination Search for Focal Pair: An interesting mixed strategy is appropriate when users have one actor of interest, and they are trying to discover other similar actors. In this case, users can use direct search to find the first actor, and similarity search to find the second actor, selecting from a ranked list of similar actors for actor 2. C-Group can be easily extended for supporting user defined domain-specific similarity functions to enable users to do a direct search for actor 1, and then use relational information (e.g. *advisor-of* or *mother-of*), to select the second actor.

5.2 Group Construction

The group context viewer (top right corner of Figure 1) shows a node link diagram which highlights the common relationships between the focal pair, and is used to visualize the dynamic changes in the relationships over time. It shows the selected focal pair as square nodes in fixed positions in the window. The focal pairs’ shared groups are shown in the middle between the pair, and their non-shared groups are shown on either side. This provides a stable, meaningful layout for a small portion of the overall network. We have found this ability to focus on a relevant sub-network particularly effective for analyzing a pair of actors [5]. The group context viewer is designed to support two major capabilities: the ability to flexibly define groups in dynamic affiliation networks and support for appropriate visualization methods which can highlight the similarity and differences in the pairs’ evolving social network.

In C-GROUP, a group is always defined as a collection of actors. C-GROUP provides a flexible mechanism for defining the groups of actors, based on actor attributes, event attributes, or participation attributes. For example, in the publication domain, users can define groups based on the institution of an author (shown in Figure 4(a)). Recall that groups defined via actor attributes will be non-overlapping. On the other hand, if users define groups based on event attributes, e.g. groups based on the topic of the paper an author has co-authored (shown in Figure 4(b)), the groups will be overlapping since the same author can publish on multiple topics. If an author in a group is clicked, all the nodes referring to the author are highlighted at once. The group context viewer is designed to provide a pull-down *Group*

By menu which lists all of the author attributes and paper attributes. Any of these attributes can be selected to define the groups.

In addition, C-GROUP is designed to make it easy to define group membership based on any temporal attributes of either the actor or the event. The group context viewer provides a second pull-down menu, the temporal attribute selector, which allows users to select from any of the attributes in the dataset having a date time type. Users can add a time window to each group, such that an actor is a member of the group only if the time attribute for the chosen attribute falls within that window. The width of the time window is set using the time window slider, which is next to the temporal attribute selector.

5.3 Group Visualization

The goal of C-GROUP is to help users understand the relationship between the focal pair and the groups defined above. In particular, we are interested in understanding the *shared* groups to which both focal pair actors are related and the *non-shared* groups (the groups in which only one of the two actors belongs), and how these group relationships evolve over time. C-GROUP provides two ways of visualizing the changing group memberships over time. One is the fixed entity layout view and the other is the dynamic group layout view. Both show changing connections over time; however each is somewhat better suited to different datasets and analytic tasks. For both views, users begin by selecting the time window size as described above. Users can drag the window over the time range, and can show a static picture of the focal pair-specific affiliation network during that time window. In addition, users can select the play button at the top of the dynamic context window, and the changing group membership will be animated as it changes over time.

5.3.1 Dynamic group layout view

This view focuses more on the shared patterns of group membership rather than on the specific actors involved in the groups. In the dynamic group view, a *group* belongs to one of the three regions: the shared region, if both *actor1* and *actor2* belong to the group, and the *actor1*-only non-shared region, if *actor1* is a member and *actor2* is not, and the *actor2*-only non-shared region, if *actor2* is a member and *actor1* is not. An actor can be the member of any region group depending on the relationship with the focal pair during the time window specified by time slider bar. In addition, an actor can be the member of more than one group in the same region if the actors are grouped by event attributes.

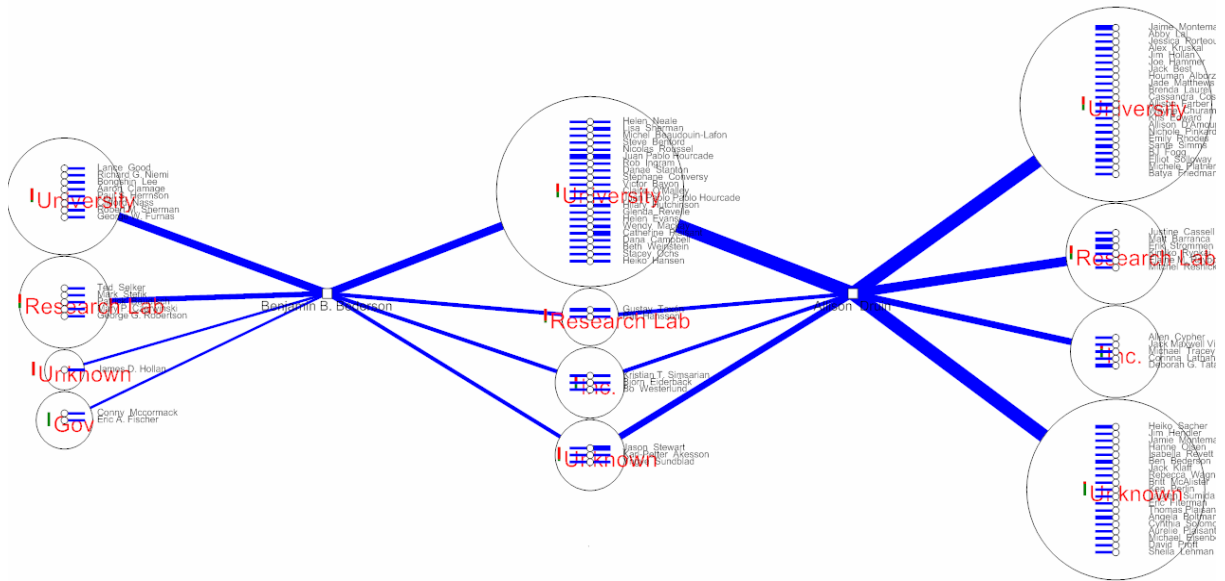
As the temporal evolution plays out, both the region that a group belongs to changes *and* the group membership changes. Users can focus only on the named groups, for example the collection of the authors that publish on a topic, as shown in Figure 4(b), or users can zoom in the groups to see the group members. Recall that an actor can belong to multiple groups, so when the groups are zoomed in, users may find an actor in multiple places.

5.3.2 Fixed entity layout view

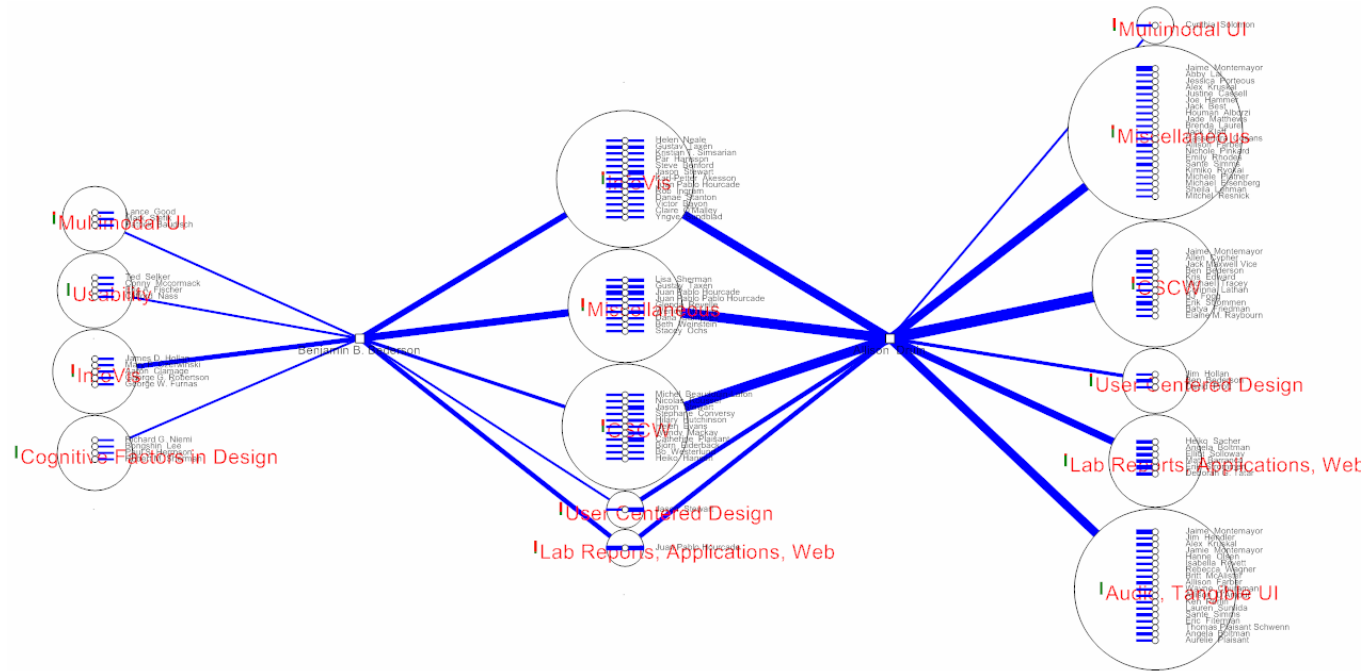
The fixed entity layout view is designed to highlight the shared relationships among the focal pair and the other actors in the domain. We consider all actors who have ever participated in an event with the focal pair. We then establish the layout of the regions based upon the overall relationship of each actor to the focal pair. The key to the fixed entity layout view is that all the entities stay in a fixed region (the *actor1*-only region, the shared

region, and the *actor2*-only region) and they are unique in the viewer. The entities may move within a region if an attribute grouping operation is used, but they do not move from shared to non-shared regions. The relationships between actor *a* and *actor1* or actor *a* and *actor2* will be indicated with edges. The entities are not duplicated even though they are grouped by event attributes. If an entity has a multi-valued attribute on the basis of an event attribute, a new group with a multi-value attribute is created instead of duplicating the entity across multiple single-valued groups (Figure 5(a)).

As the time window changes, the layout evolves. The entities are shown in each section of the viewer. The edges appear and disappear, depending on the relationship. By using this fixed entity placement layout, if the focal pair has a relationship with actor *a* independently or together, it will be visible if edges exist from the focal pair to the actor *a* during the same time window (Figure 5(b)).

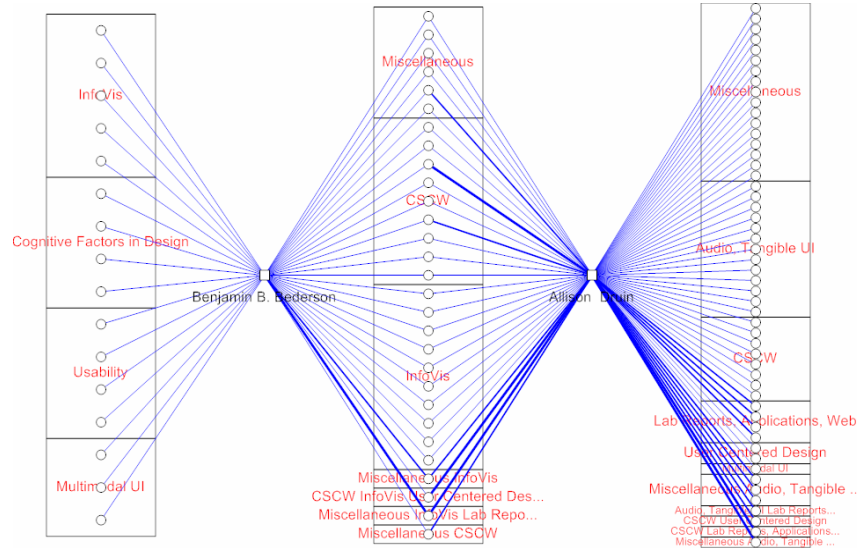


(a) Both shared and non-shared coauthors of the focal pair are grouped by their institutions(single-valued actor’s attribute). The groups are non-overlapping and an author is the member of only one group.

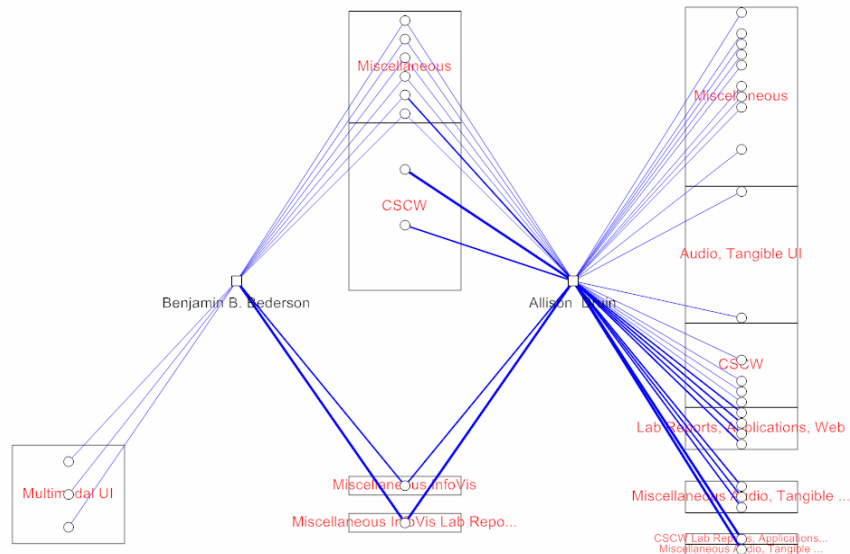


(b) Coauthors of the focal pair are grouped by topics of the papers they co-authored. The groups can be overlapping and an author can be the member of more than one group.

Figure 4. Dynamic group construction based on the attributes of actors and events



(a) All the authors (grouped by paper topics) stay in a fixed region. Each author is a member of only one group that can be defined by multi-valued attribute.



(b) The layout shows only the groups and authors of papers published between 2001 and 2002

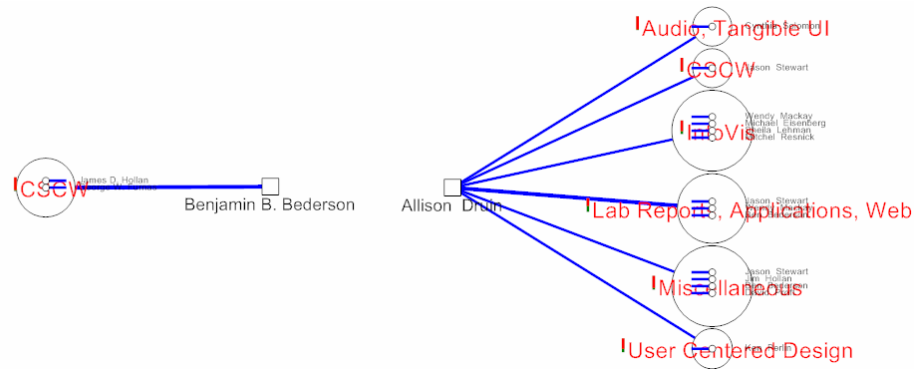
Figure 5. Fixed entity layout view

6. VISUAL ANALYTICS CASE STUDY

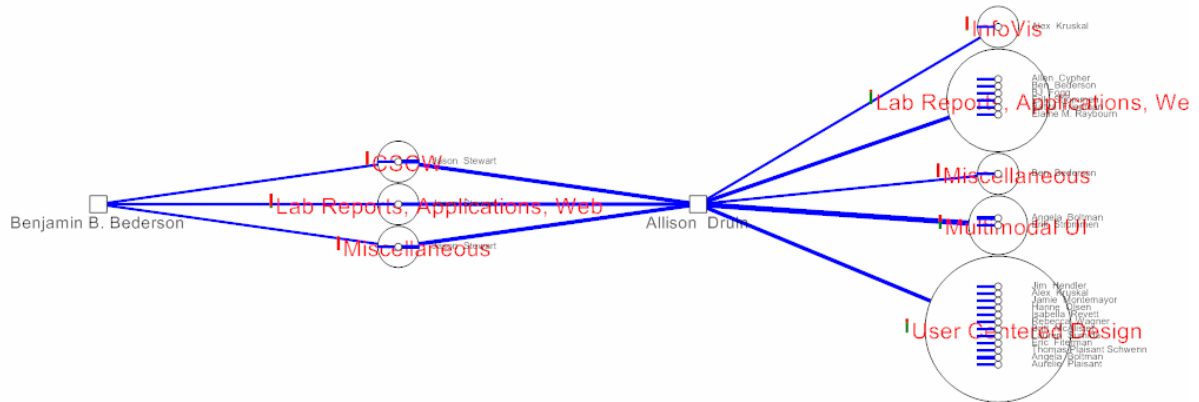
Next we present a short case study using the mentioned publication network to describe the use of C-GROUP to visualize changing group structure over time.

Drs. Druin and Bederson are professors at the University of Maryland, College Park. They are also married, making the evolution of their shared publication record interesting to analyze. For this focal analysis, we will use the dynamic group layout. Figure 6 shows the network for Drs. Druin and Bederson grouped by event attribute publication topic. We attempt to convey the animation using a series of figures taken at different times. Figure 6(a) shows the focal pair prior to arrival at University of Maryland. We can see that they initially have no shared collaborators but both publish in CSCW. Also, we notice that Dr. Druin has a larger number of collaborators at this stage compared to Dr. Bederson. The size of the nodes in the *actor2*-only regions

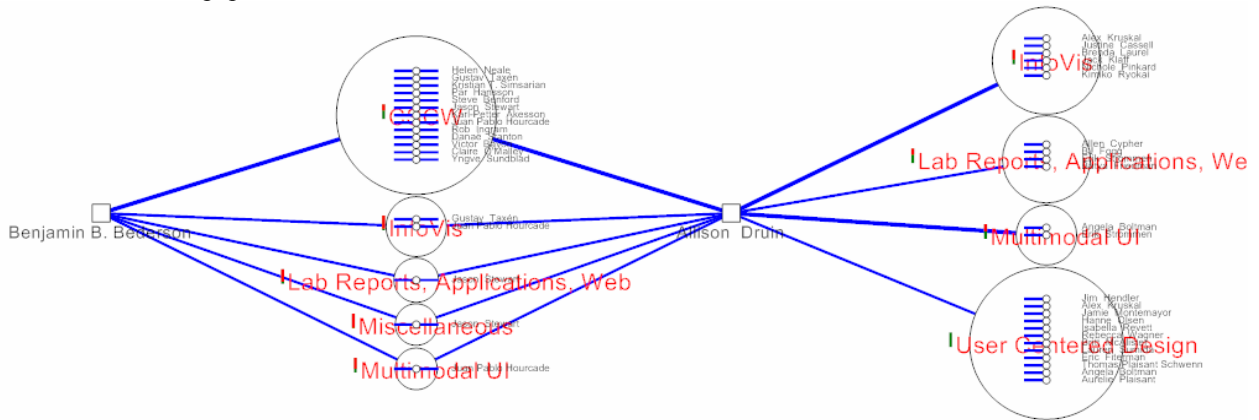
are changing because the number of authors publishing in those topics is changing over time. The next set of figures, Figure 6(b) and Figure 6(c), show the focal pair after they arrive at University of Maryland, College Park in the HCIL group. The visualization highlights a change in the network structure. While Dr. Druin's areas of publications are still similar, Dr. Bederson has expanded his publication areas; therefore, the shared area has more topics in it. Also, as time progresses, all the topics begin to enter the shared area and the size of the groups also grows. While the motion of these topics from the *actor2*-only region to the shared region is not evident from the snapshots, it is potentially insightful during actual analysis. Figure 6(d) show continued shared groups, but more individual interests in topics. Even so, the shared region continues to be the main thread of publications. Seeing this dynamic group evolution provides a different picture of the network than a static view that only shows the resulting connections over time.



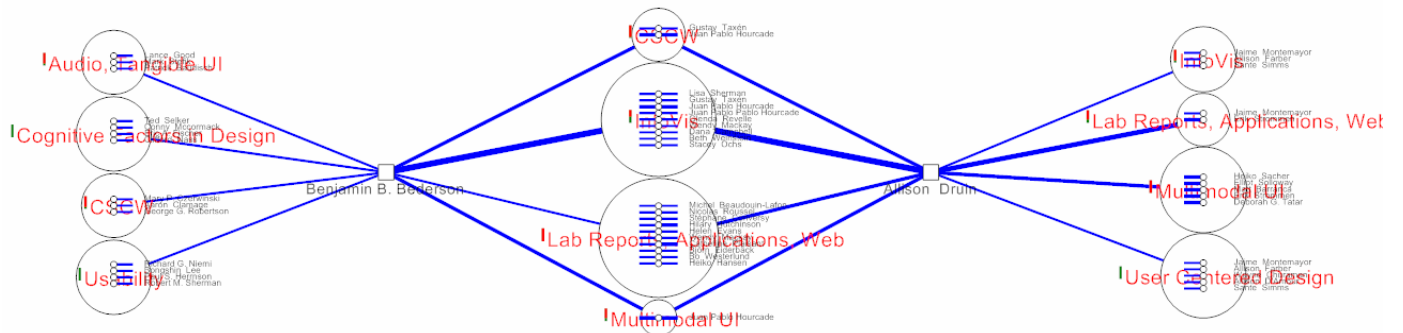
(a) Collaborators of CHI papers before 1997



(b) Collaborators of CHI papers between 1998 ~ 1999



(c) Collaborators of CHI papers between 1999 ~ 2000



(d) Collaborators of CHI papers after 2002

Figure 6. The evolution of author groups in the dynamic group layout view

7. CONCLUSION

C-GROUP allows users to visually explore dynamic affiliation networks. It allows users to select a focal pair and explore their shared relationships. Users have flexibility in defining groups based on attributes of either the actors or events in the network. Users can choose from a number of different views that highlight different aspects of the group evolution: the views can highlight the actors in the groups or the groups; the views can hold certain elements static, or animate them; the views allow users flexible control over the time granularity of the visualization; and the views support the selection of a static snapshot of a time-slice, or users can play an animation which shows the evolution of the dynamic group context. While we have chosen definitions for groups and visualizations that we found useful for our case studies, there is much room for additional research. There were several alternative definitions and designs that we considered. Some of these turned out to be ineffective, but others may prove useful in other domains, or for other tasks in the domain we have considered here. While the design of task-specific visual analytic tools for heterogeneous dynamic social networks is challenging, C-GROUP is a step toward more focused visual analysis.

8. REFERENCES

- [1] Adar, E. Guess: A language and interface for graph exploration. In SIGCHI Conference on Human Factors in Computing Systems, ACM Press, 791–800.
- [2] Batagelj, V., and Mrvar, A. Pajek - program for large network analysis. *Connections*, 21:47–57, 1998.
- [3] Baur, M., Benkert, M., Brandes, U., Cornelsen, S., M., Gaertler, K., Kopp, B., Lerner, J., and Wagner, D. Visone software for visual social network analysis. In P. Mutzel, M. Jünger, and S. Leipert, editors, *Graph Drawing Software*, Springer, 463–464, 2002.
- [4] Bender-deMoll, S., and McFarland, D. The art and science of dynamic network visualization. *Journal of Social Structure*, 7(2), 2006.
- [5] Bilgic, M., Licamele, L., Getoor, L., and Shneiderman, B. D-Dupe: An interactive tool for entity resolution in social networks. In *Visual Analytics Science and Technology (VAST06)*, 43-50, 2006.
- [6] Freeman, L. The development of social network analysis: a study in the sociology of science. Empirical Press, Vancouver, 2004.
- [7] Freeman, L.C. Visualizing social networks. *Journal of Social Structure*, 1(1), 2000.
- [8] Heer, J., and Boyd, D. Vizster: Visualizing online social networks. In *IEEE Information Visualization*, 32-39, 2005.
- [9] Heer, J., Card, S.K., and Landay, J.A. Prefuse: a toolkit for interactive information visualization. In SIGCHI Conference on Human Factors in Computing Systems, ACM Press, 421–430, 2005.
- [10] Herman, I., Melancon, G., and Marshall, M.S. Graph visualization and navigation in information visualization: A survey. *IEEE Transactions on Visualization and Computer Graphics*, 6(1):24–43, 2000.
- [11] Moody, J., McFarland, D., and Bender-deMoll, S. Dynamic network visualization. *American Journal of Sociology*, 110(4):1206–1241, 2005.
- [12] Netminer 3: Social network mining software. website: <http://www.netminer.com/NetMiner/home01.jsp>.
- [13] O'Madadhain, J., Fisher, D., Smyth, P., White, S., and Boey, Y.B. Analysis and visualization of network data using JUNG. *Journal of Statistical Software*, 1-35, 2005.
- [14] Shen, Z., Ma, K., and Eliassi-Rad, T. Visual analysis of large heterogeneous social networks by semantic and structural abstraction. *IEEE Transactions on Visualization and Computer Graphics*, 12(6), 1427-1439, 2006.
- [15] Ucinet social network analysis software. website: <http://www.analytictech.com/ucinet/ucinet.htm>.
- [16] Wasserman, S., and Faust, K. *Social network analysis: methods and applications*. Cambridge University Press, Cambridge, 1994.

About the authors:

Hyunmo Kang is a researcher in the Institute for Advanced Computer Studies at the University of Maryland, College Park. Previously, he worked for Statistical Research Division, US Census Bureau as a research scientist in 2004. He received his PhD from the Department of Computer Science of University of Maryland at College Park in 2003. His research interests are in the areas of human-computer interaction, information visualization, visual analytics, GUI design and development, and personal information management.

Lise Getoor is an assistant professor in the Computer Science Department at the University of Maryland, College Park. She received her PhD from Stanford University in 2001. Her current work includes research on link mining, statistical relational learning and representing uncertainty in structured and semi-structured data. She has published numerous articles in machine learning, data mining, database, and AI forums. She is an action editor for the *Machine Learning Journal*, is a *JAIR* associate editor, has been a member of *AAAI Executive council*, and has served on a variety of program committees including *AAAI*, *ICML*, *IJCAI*, *KDD*, *SIGMOD*, *UAI*, *VLDB*, and *WWW*.

Lisa Singh is an assistant professor in the Department of Computer Science at Georgetown University. She received her Ph.D. from Northwestern University in 1999. Prior to joining the university in 2002, Dr. Singh was a database adviser/architect and data warehouse consultant for small- to mid-size firms in the Bay Area. Current research projects include visual mining of graphs and social networks, sampling techniques on large graphs for data mining applications, data modeling and integration of observational scientific data, and data mining in the context of privacy preservation for large time series data sets and graph data sets.