Chu, K.-H., Suthers, D. D., & Rosen, D. (2012). Uncovering multi-mediated associations in socio-technical networks. Proceedings of the Hawaii International Conference on the System Sciences (HICSS-45), January 4-7, 2012, Grand Wailea, Maui, Hawaiʻi (CD-ROM). New Brunswick: Institute of Electrical and Electronics Engineers, Inc. (IEEE).

Uncovering Multi-mediated Associations in Socio-Technical Networks

Kar-Hai Chu
University of Hawaii
karhai@hawaii.edu

Daniel D. Suthers
University of Hawaii
suthers@hawaii.edu

Devan Rosen
Ithaca College
drosen@ithaca.edu

Abstract

People interact in a variety of ways, and often choose what media to use based on the relationships they have with their interlocutors. Similarly, there are different ways through which people are associated in socio-technical networks or "online communities." This paper reports a study that characterizes how the associations between members of a socio-technical network—the Tapped In network of educational professionals—are distributed across various asynchronous media, and what clusters of associations suggest about community structure within this network. The paper also illustrates an application of an analytic framework we are developing for extracting and analyzing interaction and affiliation data from log files. Affiliation networks of actors and media artifacts were constructed in which directed arcs relate actors to the artifacts they read, write or edit. Visualization of these graphs and associated sociometrics demonstrate how affiliations between participants are distributed differently across media types, reflecting the different roles these media play, and revealing community clusters within the network.

1. Introduction

More people each day are connecting with each other through various socio-technical networks. Whether for education, information seeking, friendships, work or other reasons, diverse technology-mediated affiliations are being formed. The emergence of Web 2.0 further affords the everyday user a more active and participatory role in all aspects of the online world [1], and offer many ways through which people can connect and interact with each other. Whether through public forums such as Facebook's "wall" feature, live video chat via Skype, or private messages, users today have more options in deciding what media to use for interactions. As the options continue to grow, researchers need to be able to study multiple user interactions taking place simultaneously in different media. Licoppe and Smoreda [15] claim that "… at the level of construction and experience of social ties, the nature and the role of the mediation technologies that are used to support them matter." Their research found that the choice of technologies by which people share personal news or keep in touch with each other both *reflects* and *reaffirms* the nature of the relationship between the interlocutors.

The present line of research is inspired by this idea that associations in online communities and other socio-technical networks are also distributed across media in ways that reflect and reaffirm participant relationships. Rather than assume what kinds of relationships exist, we seek to discover the full range of ways in which dyads are associated with each other via the available media. We are further inspired by the concepts of networked individualism [3, 31] and socio-technical capital [20]. Diverse forms of associations between participants, including weak ties [9] and resource sharing as well as strong socio-emotive ties are of value in a networked society. The sharing of resources benefits both the individual users within these networks and the communities formed by them [10, 31]. Therefore, we are exploring the ways and means of resource sharing and relational contact.

This paper reports an exploratory study of the Tapped In professional network of educators [6, 24] to find what kinds of artifact-mediated associations are present, and how they cluster into apparent communities. The study is based on a generalized notion of "ties" that includes not only direct interaction between individuals, but also situations in which the individuals do not directly communicate yet have some kind of association by virtue of their mutual access to artifacts. Following Latour [12], we use the term "association" to include various artifact-mediated pathways of influence between actors. We prefer "association" to "relationship" because the latter has socio-emotive connotations, while we include in our scope of consideration the diversity of other ways in which people interact with or influence each other [3, 10], or even potentially may do so in the future [25]. We treat questions of what kinds of associations exist and whether they endure as empirical questions rather than making a-priori assumptions. We prefer "socio-technical network" to "online community" because "community" is over-used for its positive connotations. Similarly, we treat questions of how associations cluster in a socio-technical network and whether these are communities as empirical questions.

This paper addresses two related research questions concerning technology-mediated associations:

- What artifact-mediated associations exist between members of the Tapped In socio-technical network?
- What does this tell us about the role technology plays in enabling associations between members?

These two questions are two sides of the same issue. There are many ways through which people can interact with each other in Tapped In, each with different affordances. We can characterize a given type of association in terms of the pattern of technology mediation in which it lives; and we can characterize the role of a given technology medium in terms of the associations it supports.

There are different ways in which researchers can study socio-technical networks. For example, one can summarize participants' activities into "ties" and use social network analysis to see what types of social structures exist and the different roles that people play [21]. At a finer granularity, the sequences of actions constituting the interactions of a pair or small group can be studied using microanalysis such as conversation analysis [18]. However, social network analysis does not give us much information in the details of a tie, while microanalysis does not provide a global view to situate the interaction being studied in a wider context [23]. We propose a meso-level analysis of how people interact via different forms of media. This middle level can facilitate moving between the social network and sequential interaction granularities of analysis [28], but this paper shows how it is also useful in its own right. In the study reported here, we analyze associations between members of a socio-technical system that form by virtue of their mutual interaction with certain digital artifacts.

The remainder of the paper provides a brief introduction to the Tapped In community, discusses the analytic approach taken, and then steps the reader through the analyses, results obtained, interpretations, and contributions.

2. Tapped In

The study examines participant interaction in SRI International's Tapped In (tappedin.org), an international (albeit mostly US) network of educators engaged in diverse forms of informal and formal professional development and peer support [6, 24]. Tapped In has hosted the content and activities of more than 20,000 education professionals in more than 800 user-created spaces that contain threaded discussions, shared files and URLs, text chats, an event calendar, and other tools to support collaborative work. More than 50 "tenant" organizations, including education agencies and institutions of higher education, have used Tapped In to meet the needs of students and faculty with online courses, workshops, seminars, mentoring programs, and other collaborative activities. Also, approximately 40-60 community-wide activities per month were explicitly designed by Tapped In volunteer members to help connect members. The Tapped In codebase is available under an open-source (GPL) license. Extensive data collection capabilities underlying the system captured the activity of all members and groups. SRI colleagues provided eight years of data to us. Out of this data, we selected a period of peak usage that occurred from September 2005 through May 2007 for analysis in this study.

Because Tapped In is populated with members of multiple tenant organizations as well as unaffiliated members, it is best seen as a network of education professionals rather than a single "community." Members may move freely between most forms of participation. The question of what communities (or clusters) exist in this network is a matter for empirical investigation. We approach this question in terms of the artifact-mediated associations found between members—an analytic approach described in the next section.

3. Analytic Approach

After briefly reviewing social network analysis, we introduce the entity-event-contingency graph approach by which we bridge between events and interaction on the one hand and social networks on the other hand. Then we detail the intermediate level of analysis at which this paper operates: associograms.

3.1. Social Network Analysis

A social network is generally defined as a system with set of social actors and a collection of social relations that reflect how these actors are relationally tied together [30]. Network analysis serves two purposes, revealing underlying social structures and discovering interactions among social actors. These two purposes are addressed simultaneously: the system's structure is identified through examining the relations among the system components [22]. The increased use of mediated information and communication technologies (ICTs) that leave traces of interaction has increasingly allowed for the precise measurement and study of social systems.

There are different ways by which social network analysis can help us study online communities. The *sociogram* graphical representation is a powerful analysis tool, able to help researchers identify clusters, central and peripheral layouts, and other structural properties that otherwise might not be as obvious with numerical data. Different studies use sociograms to help study organizational efficiency, structural bottlenecks, and informal relationships [2, 5, 7]. *Affiliation networks* provide a different lens into online communities. Represented as bipartite graphs, these networks connect actors to the groups or events they participate in. Generalizing this idea, two actors can be thought of as affiliated if they have both accessed a mediator, such as a digital artifact. Affiliation networks differ from single mode networks in several ways. They are bimodal, or two-mode networks, where a set of actors and a set of events each constitute a mode. The graph is bipartite in that links go from members of one mode to members of the other. Connections among members of one of the modes (e.g., simple ties between pairs of actors) are based upon the linkages formed via the second mode. One of the benefits of the bipartite approach is that the composition of bimodal networks can be represented simply and explicitly in a single framework.

3.2. Entity-Event-Contingency Graphs

Activities in online communities are often distributed across multiple media and sites. As a result, traces of such activity may be fragmented across multiple logs, disassociating actions that for participants were part of a single unified activity. Also, the granularity at which events are recorded may not match analytic needs. For example, media-level events are the wrong ontology for analyses that begin with interactions and ties rather than isolated acts. Translation from log file representations to other levels of description may be required to begin the primary analysis. Since interaction is distributed across space, time, and media, and the data comes in a variety of formats, there is no single transcript to inspect and share, and the available data representations may not make interaction and its consequences apparent.

To address these concerns, we have developed an abstract transcript representation called the Entity-Event-Contingency (EEC) graph that provides a unified analytic artifact [27]; and an analytic hierarchy derived from the EEC that supports multiple levels of analysis [28, 26]. To construct the EEC, log files are abstracted to collections of events (actors taking actions on media objects), supported by a domain model describing the participating entities (actors, actions and digital objects or "artifacts"). Other information is also recorded, such as time and virtual location. To support sequential analysis of interaction, directed graphs that record observed relationships between events called *contingencies* are constructed [27]. The present study does not use contingency graphs, but rather abstracts further to *associograms*, two-mode directed graphs that record how associations between actors are mediated by their creation and modification of and access to digital artifacts.

3.3. Associograms

Associograms are like affiliation networks, but they relate actors to mediating artifacts and can be directed: arcs are directed from actors to the artifacts they read, and from artifacts to the actors who wrote or edited the artifact (Figure 1). The arcs represent a form of dependency: they may be reversed to indicate direction of information flow. Existing social network analytic techniques for affiliation network analysis may be applied to associograms, or transitive closure can be used to convert associograms to sociograms to which other existing techniques can apply [30]. The results of network analysis can then be interpreted by reference to the other levels of analysis.

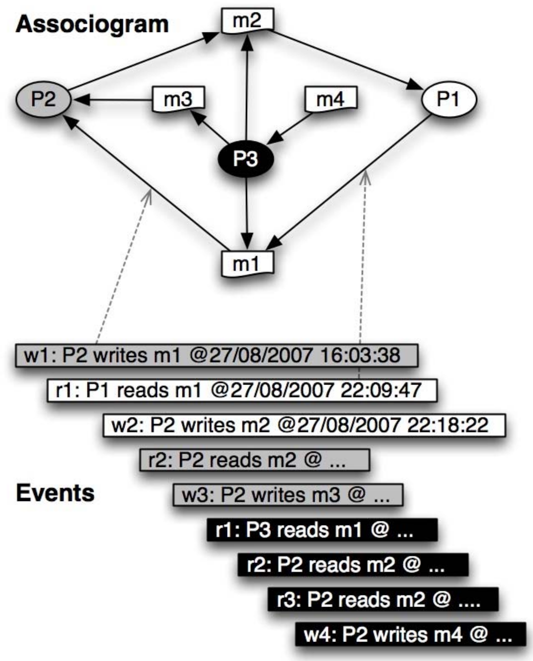

Figure 1. Constructing an associogram from events. For example, the event of P2 writing message m1 is represented as a dependency of m1 on P1, and the event of P1 reading m1 is represented as a dependency of P1 on m1. The cycle between P1 and P2 indicates an interactional relationship, while P3 is a consumer of artifacts created by P1 and P2.

Thus associograms bridge between interaction data and network analysis.

Associograms offer a different viewpoint of a network than sociograms. An associogram tie links an actor to an artifact; therefore, an artifact always mediates an association between two actors. Another way to visualize this is to expand a sociogram tie, identify the means by two actors have interacted, and then explicitly display the mediating factors. Because digital artifacts afford different types of interactions, associations between actors need not be the result of "direct" exchange or transmission of information. While the tie in a sociogram might represent a relationship (e.g. family member) or metric that defines the relationship between the actors (e.g. frequency of contact), an association defines the relationship in terms of the distribution of affiliations across a set of mediating artifacts.

Sociogram ties compress all information concerning how two actors are connected: ties are aggregated descriptions of all their interactions. While this is of great use in studying large-scale social relationships in a network, it does not provide any information about how the socio-technical components of an online community contribute to those ties. In short, it cannot answer *how* users connect with others within a socio-technical system, and whether the different media types affect the relationships and structures that are formed. Associograms are a first step to study the relationships between actors and artifacts, helping to understand the means by which actors are creating associations with each other.

In summary, we are using a unique method of analysis that examines the "middle ground" between network and interaction analysis of online communities. Social network analysis and its use of the classic sociogram generally take a high level view of user relationships, but do not provide how those relationships are formed or mediated. Various micro-analytic methods provide detailed information for different forms of interactions, but do not expose structures of the larger network. This study focuses on a middle level of analysis: associations between participants that take place through digital media.

4. Analysis

The analysis provides an initial view of how associations between participants are distributed across media. We describe how the associograms were generated from source data, and the various visualizations used in the analysis.

4.1. Generating Associograms

We used software built within the EEC framework to load and categorize the data from Tapped In. Approximately 7.8 million rows of data were imported from the database tables provided to us by SRI. The EEC describes events in terms of their participating entities: the actor and the artifact manipulated. This description was used to construct various associograms. The present analysis focused on how users are associated via four different asynchronous artifact types: files, discussions, notes, and calendar entries. Future publications will report our ongoing analysis of voluminous data from synchronous interaction in chat forums.

For each event retrieved from the EEC data, we identified the actor and artifact involved, and created a tie from the artifact to the actor. In this way, we created simple, albeit quite large, VNA formatted files. VNA files are basic text files used by NetDraw[1], a software package that can visualize networks and their associated sociometrics. VNA files include the core network data as node and edge lists, and also include node attributes, allowing for the visualization of attributes (e.g., via the color or size of the node). We identified each node as either an actor or artifact. Each of the VNA files was imported into UCINet[2] and NetDraw. After the VNA files were loaded, the graphs were laid out using spring embedding based on geodesic distance. A useful property of this layout algorithm is that nodes with similar connectivity are clustered together. For larger networks, we switched to the Gephi[3] visualization package. Still in the alpha development stage as of this writing, Gephi does not have the stability of UCINet. However, it uses a sparse graph representation that supports visualization of larger networks than UCINet can handle. It also offers more diverse layout methods than UCINet, using newer algorithms that utilize multi-core processors, such as OpenOrd [17]. From this point, the analyses diverged as we took different views of the network data, as described in the next subsection.

4.2. Visualizing Associograms

The analysis was undertaken at two granularities: examining how actors are associated via different *types* of asynchronous artifacts (files, discussions, notes, and calendar entries); and how they are associated by specific *instances* of artifacts. At the first granularity, we began with a general view of how associations were distributed across the different Tapped In artifact types by grouping all events according to which artifact type they operated on. Figure 2 shows the resulting associogram; Blue nodes (small dots in grayscale) represent individual actors and red nodes (lighter rectangles grayscale) represent one of the four possible artifact types. Node sizes represent the graph-theoretic degree of the node. We used a spring embedding layout from UCINet for this visualization; due to its limitations in handling large networks, only a random sample of the full dataset was included.

Several analyses undertaken at the finer granularity expanded each of the four artifact types into their component individual artifacts; that is, instead of having all instances of an artifact type represented by a single node, each artifact instance is now represented by its own node. Figure 3 shows the entire network of actors and artifacts. The node sizes represent the degree of the nodes. Additionally, the actor node with the highest degree is highlighted in green (lighter node left of center in greyscale); another node is highlighted in yellow (light node below the center) to represent the highest ranked artifact. This figure does not distinguish whether the node is an actor or artifact. This was a deliberate decision to help us answer certain questions that address both actor and artifact equally. In these situations, we refer to the node as an "actant." An actant is a Latourian concept that views both actors and artifacts in the same light [12].

After viewing the network graphs that were inclusive of the four artifacts, we continued our analysis by focusing on two specific artifact types: files and discussions. Both files and discussions are common in many different online communities and have functionality that is understood by average users.

[1] http://www.analytictech.com/netdraw/netdraw.htm

[2] http://www.analytictech.com/ucinet/

[3] http://gephi.org/

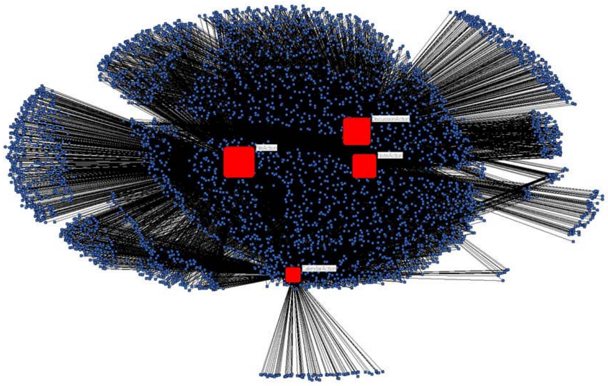

Figure 2. Associations between actors (small blue nodes) and four artifact *types* (large red nodes). Node size is degree.

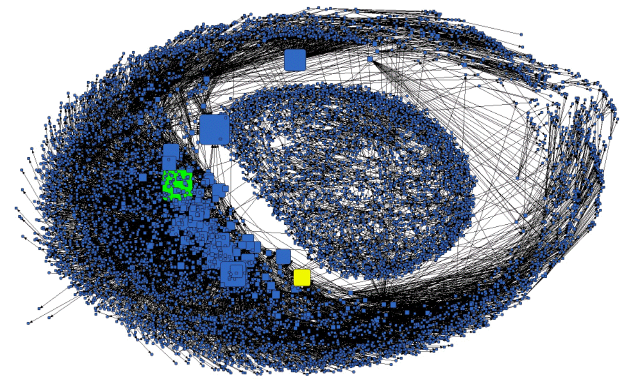

Figure 3. Associations between actors and artifact *instances.* Node size is degree; green node is highest degree actor and yellow node is highest degree artifact.

Regarding files, we considered events that included creating, editing, or downloading a file. For discussions, we included creating, viewing, or replying to a discussion posting. Figure 4 shows the associogram of files and actors, and Figure 5 contains actors and discussions. These graphs are now generated with Gephi to allow visualization of the full two-year datasets. In both cases, we separated actors from artifacts, where the red nodes denote artifacts and blue nodes represent actors. Another difference in Figure 5 is that node sizes now vary by betweenness centrality, rather than degree. Betweenness centrality is a measure of the extent to which a node lies on the shortest path between pairs of other nodes [30]. High betweenness suggests that the corresponding entity is an important mediator of the associations between other entities. The last figure (Figure 6) returns to viewing the file associogram, where node size represents betweenness.

5. Results

Figure 2 is a general view of the distribution of interactions in the Tapped In community. The associogram gave us an idea of which artifacts mediated the most activity. The graph also generated several new questions. Recall that nodes with similar connectivity are clustered together spatially. It is clear that the majority of the actor nodes are in the center cluster, revealing that most members use at least two of the four artifact types being studied. However, there are clusters at the periphery that show members who are associated only via a single artifact type (as seen by the lines extending from them to the artifact type nodes in the center). While it is not unusual that members might only use files, discussions, or notes, possibly dictated by how leaders or moderators of their Tapped In groups might lead activities, it is strange that some users only interact with calendar events. Generally, calendar event interactions are administrative duties that members can do to create, edit, view, or otherwise manage scheduled chat events that take place within Tapped In. We randomly selected several of the members that were in the sub-cluster, and found that these members did not have much activity in the community outside of viewing one or two scheduled calendar events. This was further confirmed by viewing some of the chat transcripts from the events, where some participants said that they were unfamiliar with the Tapped In community, and only followed a link to participate in a chat-based scheduled calendar event. Our current work to add chat sessions as artifacts will help fill out the picture of these participants' activities.

Figure 3 is a continuation of the analysis of Figure 2, expanding each artifact type into its individual components. Here, we treat all node types as the same; no distinction is made between whether a node represents an actor or artifact. This helped us determine whether there were more people who interacted with many artifacts or artifacts that connected many people. Sorted by degree, the thirteen largest nodes are all actors; the highest ranked node has 214 ties with artifacts (Figure 3, highlighted in green). Only at the fourteenth ranked node did we find an artifact, a file that had seventy distinct ties with actors in

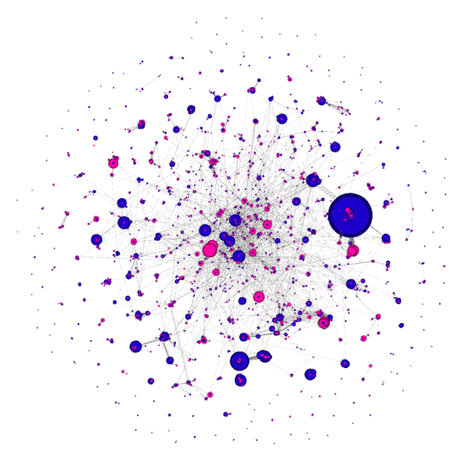

Figure 4. Associations between actors and *files only.* Node size is degree.

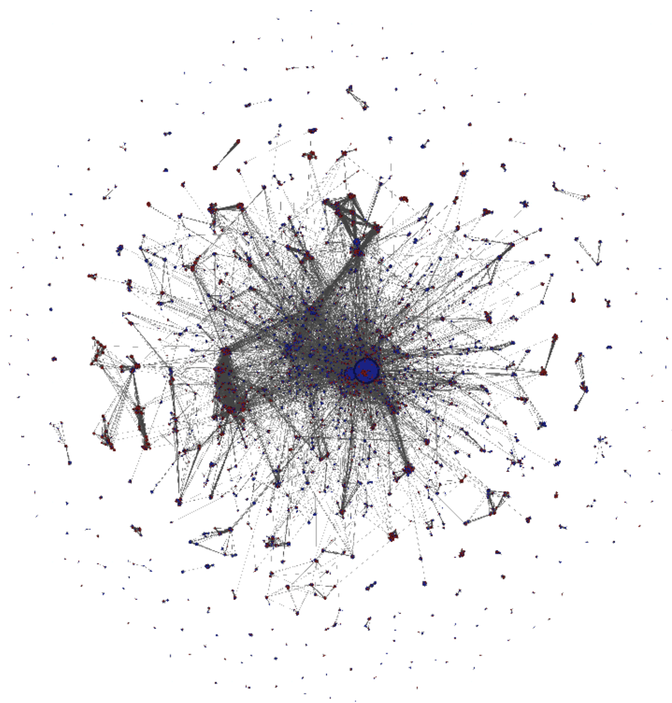

Figure 5. Associations between actors and *discussions only*. Node size is betweenness.

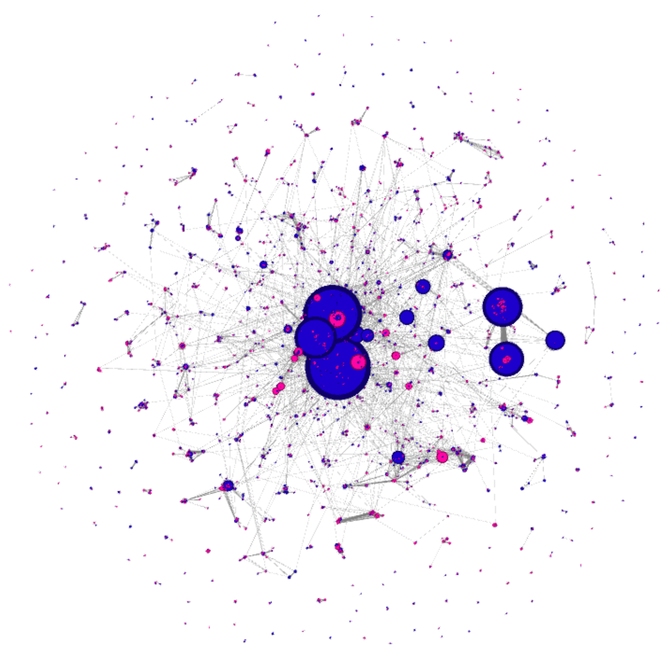

Figure 6. Associations between actors and files. Node size is betweenness.

the network (Figure 3, highlighted in yellow). The prominence of actors in the degree centrality ranking indicates that there are a set of core individuals that access many files or interact with a large number of other users, as opposed to a set of artifacts dominating the core of the network. A possible explanation for this result may be that central users are more influential than central artifacts in overall connectivity, but more exploration of the specific cases is needed to validate this claim.

Having generated Figures 1-3 with UCINet and NetDraw, we switch to Gephi for Figures 4-6 to examine networks that include actors and their associations with only a single digital artifact. Figure 4 filters the graph in Figure 3 to only include interactions with file artifacts. We also differentiated actors (blue) and artifacts (red) by color. Structurally, it is noticeably different than the previous graph containing multiple artifact types, although this can be attributed to using different layout algorithms between the two software tools. We again used node size to represent degree centrality. When ranking the nodes based on the number of ties, the top ranked node was still an actor, but in this case, a file artifact was the third highest ranked node. Compared to Figure 3, this would suggest that many of the actors with high degree centrality are associated in ways other than through files in Tapped In.

Similar to Figure 4, Figure 5 also examines a single artifact type: discussions. In this instance, we varied node sizes based on betweenness instead of degree. The significance of nodes with high betweenness is that they may enable or control movement of information in the sociotechnical system. Figure 5 shows that the nodes with the highest betweenness are all actors, indicating that a few individuals participate in discussions of otherwise weakly connected clusters, so potentially play an important role in connecting sub-communities across the network. In examining the details of the discussions with high betweenness, one of the highest ranked ones is located in Tapped In's "Reception" room, where all members are directed when they login. These discussions have a high potential of being boundary objects that can connect members from different sub-groups and offer a common virtual space to interact regarding intersecting topics.

For Figure 6, we return to files and examine their betweenness values as well. It is noticeably different from the discussion associogram; there are more highly ranked nodes, although the highest ones are mainly actors. Many of the largest nodes are also connected to each other (via discussion associations). The difference between the two networks can also be seen by the variation in the several network measurements. The average node degree for files was 5.4 and the average path length was 10.4. For discussions, the average degree was 8.4 and the average path length was 8.9. These values demonstrate that when viewing the Tapped In community under different artifact-specific lenses, there are contrasting methods in how users communicate, cluster, and interact via artifact mediation. This finding indicates that communicative interaction in this community has a greater potential to bridge unique users than the artifacts created by the users.

6. Discussion

The top-level associogram (Figure 2) was a good start to see how activity was distributed across different artifacts, and how the online community interacts as a whole. It also helped us identify a potential gateway through which guests and other new users were entering the Tapped In community. In many cases, new users were introduced to Tapped In by looking up information regarding a scheduled calendar event. From a design perspective, it could be useful to provide better navigational utility to calendar events that would help guest users discover other parts of the community. Continuing to Figure 3, we start to see whom the important actors and artifacts are. The associogram, combined with network metrics, helps us to identify who is interacting with many different artifacts, and what artifacts are connecting the most people. However, both of these graphs still view the community at a high level, and only by zooming further into individual artifact types do we begin to find how the different artifacts vary in their mediation.

When we examined the file associogram (Figure 4), we expected to see certain members on the top of the list of high degree or betweenness. Tapped In has a core group of power users that help support the different communities via content contributions, helping and guiding new members, organizing online events, as well as in many other ways. They come from a variety of backgrounds, whether teachers, educational professionals, system administrators, SRI researchers, etc. Surprisingly, the highest degree actors were standard users with different backgrounds, including a university instructor, a doctoral student, and an official at an education center. In contrast, the betweenness graph for files (Figure 6), shows that many of the highest ranked members are the power users. This is to be expected, as many of these users connect different sub-communities as part of their activities when helping new members or organizing different events. However, one regular member (RM) was found in the top ten under both measurements. Under a casual glance, it appeared that this resulted from RM creating many hundreds of files, thus being the only path between any two given files they created. However, under closer scrutiny, we can see that many other users download RM's files. These other users appear in two distinct sub-groups, suggesting that RM was a leader and a content contributor who employed file sharing as an important part of their interaction with other users. An even closer examination of the patterns of user activities could help us validate our hypothesis, but the combination of centrality metrics and network structures were able to help us pinpoint this potential group for study.

The discussion associogram offers both information regarding the interactions in the community and direction to where additional analysis is needed. Visually, the discussion network is noticeably denser than the file network. This structural variance is confirmed by the metrics for average density and path length. Even though both networks display many small sub-groups at their respective peripheries, the core of the discussion network is very well connected when compared to the file network. This suggests a much higher level of inter-group communication via discussion forums than shared files. We have further evidence of this by comparing the betweenness graphs of the two artifact types (Figure 5 and Figure 6). The discussion graph only has two significantly large nodes, both actors situated near the center of the network. As both actors are Tapped In power users, this result is not unexpected. However, in additional to the highest ranked power users, the file network also contains several more large actor nodes, most of who are not power users. These are all potential boundary spanners that are connecting different groups via shared files. This indicates that the community as mediated by files remains in more separated clusters, and in many cases, is only interconnected through a few boundary spanners that participate in multiple sub-groups. Within the discussion network, there are very few visually identifiable actor or artifact nodes with high betweenness, demonstrating that it is more likely that users will participate in discussions outside of their natural groups. The significance of this finding is two fold. First, it indicates that different sub-communities interact with each other and close structural holes [2] via different forms of boundary spanning, e.g. discussions *and* artifacts. Second, that the analytic approach used can indeed uncover these boundary-spanning examples. We identified one of the discussions with the highest betweenness values to reside in Tapped In's "Reception" area, which might be a focal point of boundary spanning activity. The reception area is accessible by all users, and contains many resources that help members with how to use and navigate the Tapped In system.

The difference in the associograms for files and discussions suggest that the types of interactions being mediated by the two artifact types are different. "Important" files, as defined by their high betweenness and degree measures, all appear to mediate interactions between actors with similarly high betweenness and degree. This sub-community of power users and artifacts suggest that files are not widely used in the Tapped In community, but rather limited to experienced users who are more familiar with the system, and shared only within sub-communities that users belong to. The denser discussion associogram, combined with few high betweenness nodes, indicate that people are more likely to interact via discussions in different communities, and that discussions mediate interactions of people from different communities, and are potential boundary objects. A next step would be to complete a more qualitative analysis on the content of the discussions. This would help determine whether the discussion topics and conversations are driving boundary spanning activity, or whether it is another factor, such as the virtual rooms that the discussions reside in. Files appear to mediate a different type of interaction. The associogram shows that many of the high degree nodes also have high betweenness values, which is a sign that there are concentrated intra-cluster interactions. While ties still exist to connect different sub-clusters, they are significantly less than ones that were in the discussion graph. While both artifact types can be accessed in similar ways, it is likely that the affordances of each mediate different behavior in how they are used to connect with others.

The results are relevant to research and design of online communities. First, the EEC framework has proven to be a useful analysis tool. The associations that it creates helped us identify how digital media differ in how they mediate user interactions in an online community. The associograms enabled us to compare different artifacts and revealed significant structures via a single artifact type, and helped us obtain a better understanding of how people use files and discussions differently. Online community designers who are attempting to increase boundary spanning between different sub-communities can make better use of discussion forums to promote the activity. Calendars of events might be leveraged as a gateway to the community and connected to tools and activities to keep new users in the community. Making discussions visible network-wide can also help users from different groups find each other and potentially create new interactions. File artifacts can compliment the use of discussions by offering users a means of intra-community communication and knowledge distribution. Lightweight interaction tools associated with files may help improve intra-community interaction. Designers should not assume that online community members will make serendipitous discoveries of files they were not looking for, but rather design for such serendipity.

7. Contributions

Although prior research has analyzed the structure of online communities [8, 13, 29], we were seeking an analytic approach that can flexibly move between different modes of communication and across various interactive platforms. As such, contributions from this research are both conceptual and operational. A theoretical framework was needed that could provide an abstraction of the data appropriate for analyses to bridge across ontologies, and the EEC framework was used to conceptualize and generate the network data and related findings presented above. One of the unique contributions of the current research is that the operationalization of network data from the EEC illuminated community and artifact structures that have remained cloaked in previous analyses of online socio-technical networks. For example, we report the finding that different forms of digital artifacts and modes of interaction mediate different types of users in unique ways. It is important for us to point out that our research agenda was not just about looking at “the” structure of a particular online community and comparing it to others, but rather is showing that one can find different structures via different media and raising questions concerning implications. If different structures are found in different media, might there be multiple communities, each distributed differentially across the media? Although we posit that questions arise from both our findings as well as our theoretical approach, the window gained from our analyses proves a view from which we can start investigating some of these questions.

8. Future Research

There are several directions that future research on this area can address. In our current study, associogram ties represent associations between an actor and a single artifact type. We plan to follow up our study with additional artifact types, such as synchronous chats, to see how they compare against files and discussions. Challenges in this work have included the size of the data and defining what constitutes an “artifact” in this spatio-temporally structured medium: the room within which chat takes place? Scheduled events only (and do recurring topics count as one or different events), or also serendipitous interactions? Additionally, instead of taking a comparative approach, we would like to see how interactions between two actors are distributed across multiple artifact types taken together, and how this multi-mediated association reflects the relationship between actors. Relationships between actors could be characterized by common patterns of distribution across media [15]. Direction of association will be important in such analyses. Whether through weak ties [9] or strong ties [11], the relationships that we have with others are based on a number of different properties. Various indicators help identify how we are connected to other people [16], but they can be incomplete and extrapolating tie strength research into today’s virtual world has been limited. More current research [19, 14] has only recently begun exploring tie strength in different online communities. Because interactions mediated by different virtual media can be different than our we interact outside of the digital world, a more comprehensive understanding of the relationships being developed in today’s online communities could require a reconceptualization of tie strength measurement [4]. The reconceptualization of dyadic ties has implications for the community level of analysis as well. Community detection that is based on aggregate analysis of dyadic ties makes it possible to detect differential distribution of communities across different media. Further research can investigate what patterns in multiple artifact-mediated associations tell us about the community structure of Tapped In. Continued work on these problems will help further bridge the gap between low and high level methods of analysis.

9. Acknowledgements

This work was supported by NSF Award #0943147 . The view expressed herein do not necessarily represent the views of NSF. The authors thank SRI and Patti Schank for providing us with access to the data.

10. References

[1] D. Beer and R. Burrows, *Sociology and, of and in Web 2.0: Some initial considerations*, Sociological Research Online, 12 (2007), pp. 17.

[2] R. S. Burt, *Structural Holes versus Network Closure as Social Capital*, in R. S. Burt, N. Lin and K. Cook, eds., *Social Capital: Theory and Research*, Aldine de Gruyter, 2001, pp. 31-56.

[3] M. Castells, *The Internet Galaxy: Reflections on the Internet, Business, and Society*, Oxford University Press, 2001.

[4] K.-H. Chu and D. Rosen, *The utility of communication network ties*, *Proceedings of the Hawaii International Conference on the System Sciences (HICSS-44), January 4-7, 2011, Kauai, Hawai‘i (CD-ROM)*, Institute of Electrical and Electronics Engineers, Inc. (IEEE), New Brunswick, 2011.

[5] R. Cross, A. Parker and S. Borgatti, *A bird's-eye view: Using social network analysis to improve knowledge creation and sharing*, Knowledge Directions, 2 (2000), pp. 48-61.

[6] U. Farooq, P. Schank, A. Harris, J. Fusco and M. Schlager, *Sustaining a community computing infrastructure for online teacher professional development: A Case Study of Designing Tapped In*, Computer Supported Cooperative Work, 16 (2007), pp. 397-429.

[7] N. Friedkin, *A test of structural features of Granovetter's strength of weak ties theory*, Social Networks, 2 (1980), pp. 411-422.

[8] E. Gleave, H. T. Welser, T. M. Lento and M. A. Smith, *A conceptual and operational definition of 'social role' in online community*, *Proceedings of the 42nd Hawai‘i International Conference on the System Sciences (HICSS-42), January 5-8, 2009, Waikoloa, Hawai‘i (CD-ROM)*, Institute of Electrical and Electronics Engineers, Inc. (IEEE), New Brunswick, 2009.

[9] M. S. Granovetter, *The strength of weak ties*, American Journal of Sociology, 78 (1973), pp. 1360-1380.

[10] C. Haythornthwaite, *Crowds and communities: Light and heavyweight models of peer production*, *Proceedings of the 42nd Hawai‘i International Conference on the System Sciences (HICSS-42), January 5-8, 2009, Waikoloa, Hawai‘i (CD-ROM)*, Institute of Electrical and Electronics Engineers, Inc. (IEEE), New Brunswick, 2009.

[11] D. Krackhardt, *The Strength of Strong Ties: The Importance of Philos in Organizations*, in N. Nohria and R. Eccles, eds., *Networks and Organizations: Structure, Form,and Action*, Harvard Business School Press, 1992, pp. 216-239.

[12] B. Latour, *Reassembing the Social: An Introduction to Actor-Network-Theory*, Oxford University Press, New York, 2005.

[13] J. Leskovec, K. J. Lang, A. Dasgupta and M. W. Mahoney, *Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters*, Internet Mathematics, 6 (2009), pp. 29-123.

[14] K. Lewis, J. Kaufman, M. Gonzalez, A. Wimmer and N. Christakis, *Tastes, ties, and time: A new social network dataset using Facebook.com*, Social Networks, In Press, Accepted Manuscript (2008).

[15] C. Licoppe and Z. Smoreda, *Are social networks technologically embedded? How networks are changing today with changes in communication technology*, Social Networks, 27 (2005), pp. 317-335.

[16] P. Marsden and K. Campbell, *Measuring Tie Strength*, Social Forces, 63 (1984), pp. 482-501.

[17] S. Martin and W. M. Brown, *OpenOrd: An Open-Source Toolbox for Large Graph Layout*, *SPIE Conference on Visualization and Data Analysis (VDA)*, 2011.

[18] J. M. Mazur, *Conversation analysis for educational technologists: Theoretical and methodological issues for researching the structures, processes, and meaning of on-line talk*, in D. Jonassen, ed., *Handbook of Research for Educational Communications & Technologies*, Lawrence Erlbaum, Mahwah, NJ, 2004, pp. 1138-1178.

[19] A. Petroczi, T. Nepusz and F. Bazso, *Measuring tie-strength in virtual social networks*, Connections, 27 (2007), pp. 39-52.

[20] P. Resnick, *Beyond Bowling Together: SocioTechnical Capital*, in J. M. Carroll, ed., *Human-Computer Interaction in the New Millennium*, ACM Press, Upper Saddle River, NJ, 2002, pp. 647-672.

[21] E. M. Rogers and D. L. Kincaid, *Communication networks: Toward a new paradigm for research*, Free Press, New York, 1981.

[22] E. M. Rogers and D. L. Kincard, *Communication networks: Toward a new paradigm for research*, Free Press, New York, 1981.

[23] M. Schlager, U. Farooq, J. Fusco, P. Schank and N. Dwyer, *Analyzing online social networking in professional learning communities: Cyber networks require cyber-research tools*, Journal of Technology Education, 60 (2009), pp. 86-100.

[24] M. Schlager, J. Fusco and P. Schank, *Evolution of an Online Education Community of Practice*, in K. Renninger and W. Shumar, eds., Cambridge University Press, Building Virtual Communities, 2002, pp. 129-158.

[25] M. Smith, C. Giraud-Carrier and N. Purser, *Implicit infinity networks and social capital*, Information Technology Management, 10 (2009), pp. 123-134.

[26] D. D. Suthers, *Interaction, mediation, and ties: An analytic hierarchy for socio-technical systems*, *Proceedings of the Hawaii International Conference on the System Sciences (HICSS-44), January 4-7, 2011, Kauai, Hawai'i (CD-ROM)*, Institute of Electrical and Electronics Engineers, Inc. (IEEE), New Brunswick, 2011, pp. CD-Rom.

[27] D. D. Suthers, N. Dwyer, R. Medina and R. Vatrapu, *A framework for conceptualizing, representing, and analyzing distributed interaction*, International Journal of Computer Supported Collaborative Learning, 5 (2010), pp. 5-42.

[28] D. D. Suthers and D. Rosen, *A unified framework for multi-level analysis of distributed learning Proceedings of the First International Conference on Learning Analytics & Knowledge, Banff, Alberta, February 27-March 1, 2011*, 2011.

[29] J. Turner and R. Kraut, eds., *CSCW '92: Sharing Perspectives*, Association for Computing Machinery, Toronto, Canada, 1992.

[30] S. Wasserman and K. Faust, *Social Network Analysis: Methods and Applications*, Cambridge University Press, New York, 1994.

[31] B. Wellman, A. Quan-Haase, J. Boase, W. Chen, K. Hampton, I. Diaz and K. Miyata, *The social affordances of the internet for networked individualism*, Journal of Computer-Mediated Communication, 8 (2003).