UNDERSTANDING MEDICINE 2.0
Social Network Analysis and the VECoN System

Sam Stewart and Syed Sibte Raza Abidi
NICHE Research Group, Dalhousie University, 6050 University Ave., Halifax, NS, Canada

Keywords: Web 2.0, Social network analysis, Data visualization, Knowledge translation.

Abstract: Web 2.0 provides new and valuable tools to the world of medicine, and Social Network Analysis (SNA) can provide insight into how these communication networks function. This paper explores the potential for SNA methods to explain the communication patterns in the Pediatric Pain Mailing List, including identifying content experts and isolating potential subgroups of interest. These results are incorporated into *VECoN*, a novel network visualization designed to improve the standard network exploration process by presenting the network graphically and incorporating SNA statistics into the presentation.

1 INTRODUCTION

Experiential healthcare knowledge manifests in a variety of modalities: clinical case studies, problem-based discussions between clinicians, experience-based insights, diagnostic heuristics et cetera. This knowledge accounts for the intrinsic experiential know-how, insights, judgements and problem-solving strategies of healthcare practitioners. Such knowledge is not 'published' as evidence-based, yet it holds vital insights into solving atypical clinical problems. The key issues related to experiential healthcare knowledge are: (a) how to formulate a community of healthcare practitioners; (b) how to explicate and share their experiential healthcare knowledge; and (c) how to put value on experiential healthcare knowledge, especially for clinical decision making, since it is not systematically evaluated in the same manner as evidence-based studies.

In the realm of Web 2.0, the emergence of Medicine/Health 2.0 presents 'virtual' community-driven environments to create and share healthcare knowledge. The key idea is that the community creates and validates experiential knowledge in an organic manner, and applies it to provide feedback to its effectiveness. Web 2.0 based knowledge sharing mediums include online discussion forums, email-based mailing lists, web blogs, et cetera. Through these mediums, healthcare practitioners are able to articulate and share clinical, operational and even psycho-social experiences, along with insights and know-how about particular healthcare topics. The efficacy of this approach is that healthcare practitioners, originating from different backgrounds and expertise levels, can engage, collaborate and share their experiential knowledge for the benefit of the entire community.

Given the virtual nature of the community it is of interest to get insights about the knowledge sharing dynamics (active participants, key contributors of knowledge, topics of interest, influential members, et cetera) of the virtual community, as it helps to put a value on the knowledge created and shared there. We argue that a study of the community's communication patterns and of the knowledge content being shared can provide insights into experiential knowledge sharing dynamics of a specialized healthcare practitioner community. Social Network Analysis (SNA) allows us to analyze the communication networks within a socially-connected community (such as an online healthcare practitioner community) and highlight the key actors, interest groups, sub-networks, content seekers and experts, collaboration opportunities, communication barriers, and other network attributes. SNA focuses on analyzing the attributes that arise out of the structural properties of a social network, rather than the properties of the actors themselves, providing an overview of the community and how people communicate within it.

In order to make the results of the SNA useful to the PPML community at large, a visualization tool has been developed to allow the members to visually navigate the network and explore both the conversations on the mailing list and the social connections

Stewart S. and Sibte Raza Abidi S..
UNDERSTANDING MEDICINE 2.0 - Social Network Analysis and the VECoN System.
DOI: 10.5220/0003167100700079
In *Proceedings of the International Conference on Health Informatics* (HEALTHINF-2011), pages 70-79
ISBN: 978-989-8425-34-8

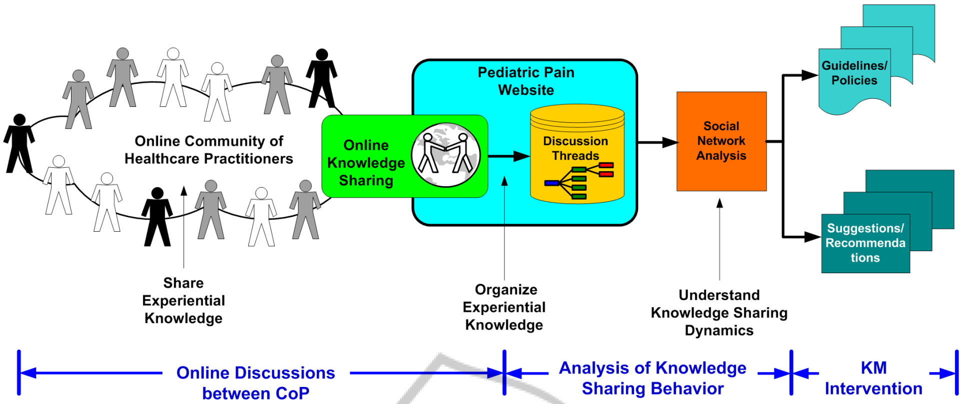


Figure 1: Depiction of the experiential knowledge sharing framework and how SNA can leverage the knowledge sharing to inform knowledge creation.

that have arisen out of it. VECoN (Visual Exploration of Communication Networks) is a tool that provides the means to visualize the social aspects of the network, along with functions to navigate to individual threads on the mailing list. A beta release of VECoN has been produced as a proof of concept of how we can provide further insights into the different uses of Web 2.0 technologies in healthcare.

2 EXPERIENTIAL HEALTHCARE KNOWLEDGE SHARING FRAMEWORK

From a health knowledge management perspective, an online mailing list or discussion forum provides a collaborative learning environment in which domain experts can disseminate their wealth of knowledge and experience, and junior health practitioners can learn by leveraging the explicated experiential knowledge. This brings into relief an experiential knowledge sharing framework (Figure 1) that allows a community of healthcare practitioners to interact and collaborate to create and share experiential knowledge, while organizing the knowledge in terms of domain specific topics. Each topic is pursed by a group of practitioners who interact through online tools, such as email or discussion forums. The discussion around a topic can be organized in terms of a 'discussion thread': a series of emails/posts by healthcare practitioners around a specific topic. Using this framework analysis of emergent social networks within the community can be performed (where the social network depicts the collaboration/communication dynamics of the community) and analysis of the content of the discussions can be done through the use of intelligent text analytics and text/knowledge creation methods.

It is anticipated that subsequent SNA of the knowledge sharing behaviour of the virtual community will provide useful insights into the patterns of knowledge flow amongst healthcare practitioners. The resultant social network will provide an objective measure of the effectiveness of the online knowledge sharing medium to support collaborative learning.

The objective of the SNA is to provide meaningful insight into the flow of communication within the network. This will be provided at both the micro and macro level. Centrality measures provide insight into the roles individuals play within the network, and will help identify content experts, those members that are actively communicating with their peers and facilitating knowledge transfer. Clique and structural equivalence analysis is performed at the macro level, identifying subgroups of actors that are tightly connected. The presence of subgroups may represent particular sub-topics of interest, or groups of clinicians that are not fully communicating with the rest of the network.

The social network analysis and visualization framework is outlined in Figure 2. The following sections will outlined the process in detail.

3 BACKGROUND

3.1 Social Network Analysis

Social Network Analysis (SNA) is the analysis of the relations between actors, i.e. understanding the underlying social structure of a community of practice. SNA utilizes the principles of graph theory from the world of mathematics to represent communication networks in terms of actors (nodes) and ties between actors (edges) (Wasserman and Faust, 1994) (Hanneman and Riddle., 2005). Traditional statistical anal-

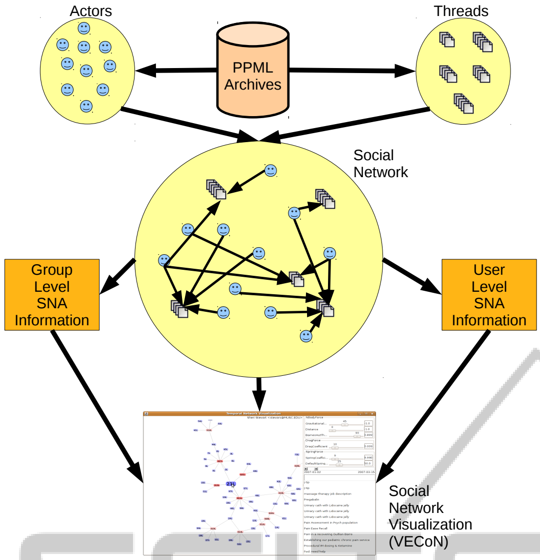


Figure 2: SNA and visualization process. The network information is extracted from the PPML database, where it is processed to produce the network. The network is used to general social networking information, which is incorporated with the network to produce the VECoN system.

ysis focuses on actors as independent units, and analyzes them in terms of their personal attributes. SNA instead focuses on the structures that emerge out of the relations between actors, and not on the actors themselves.

One aspect of SNA that is important to this project is the difference between one and 2 mode networks. In one-mode networks the nodes in the network are homogeneous, i.e., they all belong to the same class. This is the traditional network layout, in which nodes represent people and ties represent some sort of social construct that connects them: friendship, advice, work, et cetera. Two-mode networks contain two different classes of network, and ties exist only from one mode to another. These most commonly occur when one set of nodes represents people, and the second set represents events, and ties go strictly from one mode to another (indicating that a person has attended an event). A two-mode network that represents people and the events they attend is sometimes called an *affiliation* network, and is the structure of the VECoN system. The network represented in this project presents both healthcare practitioners and discussion topics as two independent classes of nodes, and the ties from a practitioner to a topic indicates that that user has communicated on the connected topic.

Because the majority of SNA methods involve one-mode networks, a common component of two-mode network analysis is to transform the data into a one-mode network. A new network is created of the members of the network, and a tie is created between them if they both communicate on the same thread. It is preferable to analyze the two-mode network when possible, as some information is lost in the transformation, but due to the existing body of literature for SNA it is necessary to perform much of the analysis on the one-mode network. Future research will be directed at adapting methods to two-mode networks.

This paper will begin by analyzing the network at the micro level, using centrality measures to identify the active members of the network. It will then move on to identifying potential subgroups of actors within the network using clique and structural equivalence analysis.

3.1.1 Centrality

The goal of centrality analysis is to recognize the most important actors in the network; those actors that are at the centre of the action in terms of communication between individuals. Three different centrality statistics are going to be presented; the actors identified through these measures occupy key roles in the network, and will be considered content experts.

Degree centrality is the simplest of the centrality measures, calculating the number of ties one actor has to the others. For the two-mode network this is the number of threads each actor communicates on, and for the one-mode network it represents the number of other actors that actor has communicated with.

Closeness centrality extends the idea of degree centrality beyond one step. Closeness centrality considers an actor central to the network if they can reach all the other nodes in the network in as few steps as possible. The calculation of closeness centrality differs little between one and two mode networks. If two actors are one-step away in the one mode network, they are necessarily two steps away in the two-mode network. Likewise, if two actors in the one-mode network share a common partner, then each of them must have shared a thread with that actor in the two-mode network. Though the normalization is different (Borgatti and Everett, 1997) the general ranking is the same. As such only one-mode closeness will be presented.

Betweenness centrality deems nodes central if they are hubs of information. Where closeness deems a node central if it can quickly reach other nodes, betweenness deems a node central if it is used as path between other nodes. A node has a high betweenness score if it falls in the shortest path between many other pairs of nodes. As with closeness centrality betweenness centrality does not significantly differ between one and two mode networks (Borgatti and Everett, 1997).

The next step in the SNA is to try and isolate potential subgroups within the network. With a topic as large as pediatric pain there may be evidence of subgroups in which actors are more active around a certain topic of interest. If the network were broken into groups, one would expect a lot of communication within groups, and relatively little communication between groups.

3.1.2 Structural Equivalence

The goal of structural equivalence (SE) is to identify nodes that occupy similar roles within the network. Formally, two nodes are SE if they have the same ties to all other nodes in the network. If two nodes are SE then one can replace the other without interfering with the flow of information in the network. In reality true SE is rare, so approximate SE needs to be measured. A simple measure would be to count the proportion of matching ties, or the number of tie changes required to make two nodes SE. There are several measures available, but for this project a simple count of the number of similar ties is used.

Regardless of which SE measure is used, a SE matrix is developed, which records the SE between all the actors. This matrix is used to group similar actors using a hierarchical clustering algorithm. The result is a binary tree, or *dendogram*, depicting a hierarchical ranking of similarities, as in figure 7. Cutting the tree off at a particular level results in partitions being created from the clusterings. The red blocks in figure 7 represent the cutpoint at which the clusters are created. Assigning the actors to these groups creates a *blockmodel*.

A blockmodel is a partitioning of the network into exclusive, non-overlapping groups, such that nodes within the group are approximately SE. For a blockmodel there tends to be a lot of communication within the blocks and relatively little between them. Once the optimal block model is determined the active blocks can be further investigated to determine the content that makes certain blocks unique.

3.2 Network Visualization

The visualization of networks is a key component of SNA, and as such there is a rich literature base describing methods of presenting networks visually. Linton Freeman (Freeman, 1999) documents the history of social network visualization from a sociological perspective, including theories on node layout (both information-based and algorithmic theories) along with the use of colour, size and shape to encode network information. There are many current tools for analytic network visualization, including `UCINet` (Borgatti et al., 2002) and an extension for the `R` statistical language called `statnet` (Handcock et al., 2003).

Previous work on social network visualization has also been directed towards network navigation. Examples include ContactMap (Nardi et al., 2002) for identifying community groups within email contacts, PieSpy (Mutton, 2004), which provides a real-time visualization of social networks for Internet Relay Chat (IRC) members, and Vizster (Heer and boyd, 2005), a tool for exploring the Friendster (www.friendster.com) social networking site. These tools are all designed for 1-mode networks, for example, the nodes in the Vizster program all represent users of Friendster, and the ties represent friendship links between them. In contrast, this project is visualizing a 2-mode network, where the first class of nodes represent mailing list members and the second represents threads, and the links between a node and a thread indicate that a certain list member has communicated on that thread.

The software being used to implement this project is the `prefuse` toolkit in Java (Heer et al., 2005). `Prefuse` was chosen because it provides a full Java library, and previous implementations of `prefuse`, including the Vizster program, have proven successful.

4 VeCON System

4.1 Visualization

The purpose of the visualization is to first provide a tool for visually exploring social networks, and secondly to provide some insight into the underlying social structure of the network. This section will outline the graph-theoretic layout decisions for the network, and then explain the visualization tools implemented to help the exploration of the mailing list.

4.1.1 Graph Structure

The network is laid out using a force-directed layout, in which the nodes repel one another and the edges act as “springs” that hold the nodes together. Because of this spring effect, the layout is also sometimes referred to as a “spring embedding” algorithm. `Prefuse` implements the Barnes-Hut algorithm (Barnes and Hut, 1986) which allows for real time calculation of spring-embedding forces. The algorithm is an iterative process, and following the lead of Vizster, this project chooses to not limit the number of iterations of the algorithm, resulting in a visualization in which the nodes migrate to their optimal positions but continue to move subtly. The effect is “a living or ‘breathing’ feel, connoting social energy or playfulness.”(Heer and boyd, 2005).

Two changes that need to be made to the spring embedding algorithm are i) an adaptation to two-mode networks, and ii) dealing with the problem of components (disconnected sections of the graph). The two-mode issue is addressed by Krempel (Krempel, 1999), in which he suggests fixing the second mode and allowing the first mode to vary. To this end, the algorithm is adjusted slightly: the first step is to disperse the threads evenly around the space, and then allow the actors to move according to the spring embedding algorithm. This method has proven to be effective, but a more formal evaluation of its efficiency is required.

The issue of components is addressed in Kamada and Kawai’s seminal work on force directed layouts (Kamada and Kawai, 1989). The solution is to partition the space according to the number of components in the graph, with each graph being allocated space proportional to its size (number of nodes). The forces are then calculated only on the nodes within the component. This means that components do not effect each other in terms of layout, and avoids the “drift” that is caused when disconnected components continue to push each other away (see figure 72 in (Fruchterman and Reingold, 1991)).

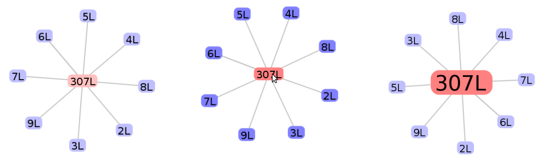
The colouring of the nodes is defined by their mode: blue indicates actors and red indicates threads. Following Vizster’s lead, nodes and their neighbours are highlighted when the mouse hovers over them: when hovering over an actor that actor’s threads are highlighted, and conversely when hovering over a thread the actors that communicated on that thread are highlighted. The highlighting is done by increasing the saturation of the colour (see figure 3).

4.1.2 Exploration

The visualization provides several different exploration methods, which will each be described in detail. The objective of each of the tools is to provide a different way of exploring the mailing list to retrieve pertinent messages.

Hover Over. As mentioned before, the visualization implements a hover-over feature. When the mouse hovers over a node, that node becomes fixed, and it and its neighbours are “highlighted”. A node is highlighted by increasing the saturation. Figure 3 demonstrates the difference between the regular and highlighted nodes.

Along with changing the representation of the



(a) Regular. (b) Highlighted. (c) Focus Node.

Figure 3: Highlighting is achieved by increasing the saturation from 25 to 50. A focus node is activated by clicking it, at which point it doubles in size.

nodes, the hover-over feature presents additional information about the node/edge in question. For actors it presents the actor’s name in the top-right corner of the visualization, and for threads it presents the thread’s subject line (see figure 8). The node names were not put on the nodes in order to avoid cluttering the visualization: current node labels are numerical identifiers, provided in order to differentiate between actors.

On the right side of the visualization is the control bar, where users can manage their search features and control the factors that effect the spring embedding. At the bottom of the control bar is a text box that presents the detailed content of the selected node. For actors a list of the threads they have communicated on is presented, and for threads the conversation is presented. When an edge is selected then that specific message is presented. The objective is to allow the user to quickly browse through the conversations, and upon finding the desired message to explore it in more detail. Future work will pursue connecting the visualization directly to the online discussion forum, but as the forum is not yet developed this feature is not implemented.

Point and Click. The objective of the point and click option is to allow the user to explore a particular node in more detail. When a node is clicked doubles in size, and the text associated with it becomes fixed. The user can then adjust the temporal filter to see where that node fits within the filtered data. While the node is in focus (i.e. while its size is doubled) the text stays static, though the rest of the highlighting effects associated with hovering remain. A second click on the node reduces it back to normal size and allows the text to change freely again. Figure 3 demonstrates what a focus node looks like.

Temporal Filtering. Adding a temporal filter provides a way to reduce the volume of information being presented. The temporal filtering is accomplished using a horizontal scroll bar, located within the control panel. The user can manipulate the lower-end or upper-end of the bar to adjust the visualization, or can set the bar to a specific window width and slide the window itself. When the mouse is released the graph is re-drawn with only those messages that are within the window presented. The user can also manipulate the dates manually by typing dates into the two date boxes that are below the bar.

When the visualization is adjusted, only those nodes that are connected to each other are presented, reducing clutter by removing obsolete nodes. There is still a potential for isolates in the filtered network, however. If an actor participates in a thread, but that contribution is not made during the window, then the edge will be removed but the actor will remain in the network, and that contribution will still be presented in the hover-over effect. Future work should explore how to incorporate the disconnected user into the force-directed layout, but for now the user is assigned a separate segment of the space.

5 METHODS

5.1 Data

This project will use the archives of the Pediatric Pain Mailing List (PPML), provided to the VECoN project by the administrators of the mailing list. There are over 13,000 messages in the PPML archives, dating back to 1991; for this project a sample of the messages from 2007 and 2008 was used. The sample has been parsed from simple ASCII text files and the messages have been written into a MySQL database. Along with sender, subject and date information about each message, a thread number has been assigned to indicate which messages combine to form conversations. The threads and senders will be used as nodes in the network, with ties between them indicating that a sender has communicated on a particular thread. Currently the PPML administrators are in the process of transferring the archives of the mailing list to an online discussion forum. Once this forum is active the VECoN system will be connected to it.

5.1.1 Social Network Visualization

Beyond simple network exploration the tool provides functions to explore the SNA dimensions of the PPML. Within the control panel there are buttons to toggle centrality indicators and block modelling. When activated the centrality indicators redefine the actor and thread nodes in the network, setting the node size relative to that node's degree. Currently only degree centrality is implemented, but all three centrality measures will eventually be added.

As well, the tool performs agglomerative clustering using the SE calculations from the blockmodelling. It is left to the user to decide how many blocks to assign to the network, presenting the blocks as "blobs" surrounding the component members. When activated the blobs modify the spring embedding algorithm to increase tensions within the blobs and decrease tension outside them, resulting in tighter groups of block members.

6 RESULTS

6.1 SNA

6.1.1 Micro Level Analysis

Table 1 lists the actors with the highest degrees in the two-mode network, and table 2 lists the same for the one-mode network.

Table 1: Actors with the highest degree centrality in the two-mode network, i.e., the actors with the most number of threads. The degree is normalized by dividing the degree measure by the maximum possible degree.

Actor	Degree	NormDegree
771	31	0.1148
901	27	0.1000
782	26	0.0963
904	22	0.0815
920	20	0.0741

Table 2: Actors with the highest degree centrality in the one-mode network, i.e., those actors that communicated with many other members of the network. The degree is normalized by dividing it by the max possible degree.

Actor	Degree	NormDegree
901	74	0.3776
904	60	0.3061
782	57	0.2908
855	56	0.2857
920	53	0.2704

From the degree centrality analysis there seem to be several actors that are quite active. The maximum normalized degree of 0.115 in the two-mode network indicates that actors are not participating in many threads, and a normalized degree of 0.378 in the one-mode network continues to demonstrate there does not seem to be a single user in the network that communicates with all other users. The histograms in

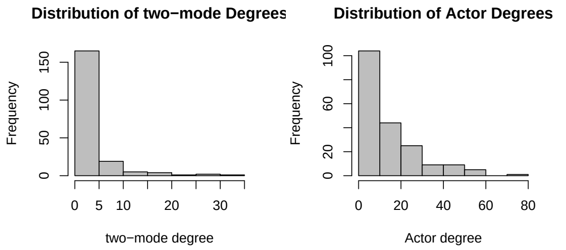


Figure 4: Distrubtion of degrees for the PPML data. The histograms indicate that there are not many actors with high degree centrality.

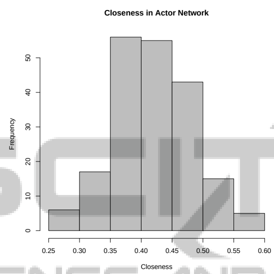


Figure 5: Distribution of closeness scores for the PPML data. The closeness scores follow a normal distribution.

figure 4 confirm that suspicion; the majority of actors in both the one- and two-mode networks have low degree centrality measures.

Table 3 lists the highest closeness scores in the network. The actors that have the highest closeness

Table 3: Closeness scores for the PPML network.

Actor	Closeness
901	0.5957
904	0.5665
920	0.5552
782	0.5521
771	0.5506

scores are the same actors with high degree scores. With a max closeness of 0.596 the network seems to have a very high closeness measure. This means that it is easy to get from any node to any other node. These findings are confirmed in the histograms in figure 5. This histogram seems to follow a normal distribution, centred at 0.419 with a standard deviation of 0.064, meaning that 97.5% of the members of the network have a closeness centrality greater than 0.292.

Table 4 lists the highest betweenness scores from the network. Once again, the same actors are prominent for the betweenness calculation. The low maximum betweenness of 0.134 indicates that the network does not depend on any particular actor to facilitate

Table 4: Betweenness scores for the actor network. As with closeness, betweenness measures the same thing in the 1 and 2 mode networks.

Actor	Betweenness
901	0.134
904	0.103
920	0.072
782	0.066
855	0.063

Figure 6: Betweenness scores for the PPML network data.

communication. The histogram in figure 6 confirms that finding, with the majority of users having low betweenness.

The centrality scores for the network indicate a healthy and active knowledge sharing network. Low degree centralities mean that there is not a single user, or a set of users, that are required to initiate conversation. High closeness scores mean that it is easy to connect one user to another, either through a thread they have both communicated on, or a short series of "mutual friends"; this is key to facilitating knowledge transfer, to allow users to connect directly to the source of experiential knowledge in as few steps as possible. Low betweenness scores compound this finding by demonstrating that, when trying to connect with someone through a series of users, there are always multiple paths available, without the need to always communicate through a single user. These results combine to describe a knowledge sharing network that has a large number of healthy communicators and an active user base.

Though the centrality results do not demonstrate the presence of any dominant users, it is noteworthy that the same actors appear at the top of all four sets of centrality measures. Table 5 lists the top ranked actors for each of the four centrality measures, and the same four actors appear in each list. If further investigation into the content of the mailing list were required, these actors should be the first to be contacted, as they seem to be the most active and most important to the communication network.

Table 5: The ranking of actors provided by each of the three centrality measures.

Degree	Degree.2M	Closeness	Betweenness
901	771	901	901
904	901	904	904
782	782	920	920
855	904	782	782
920	920	771	855

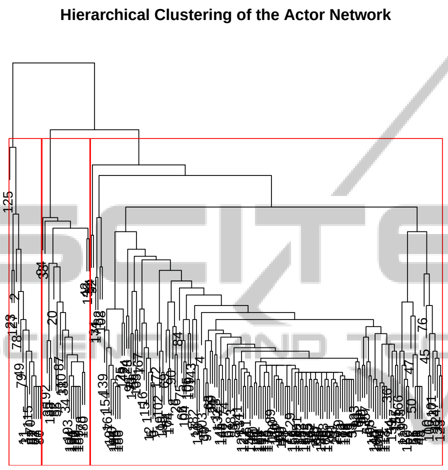


Figure 7: A hierarchical clustering of the actors in the one-mode network using structural equivalence. The red blocks indicate the points at which the tree is cut to create the blockmodel.

6.1.2 Macro Level Analysis

For the PPML the blockmodel contains three groups, two small groups with 15 and 22 members, and one large group with 160 members (figure 7 contains the dendogram). The *image matrix* produced by the block model is available in table 6. An image matrix is a table that represents the communication densities between blocks. This image matrix demonstrates that there is a lot of communication within the two smaller blocks, and a bit of communication between them (0.183 density). For the large group in block 1, however, there is little communication, either within or between blocks.

Table 6: Communication Densities Within and Between Blocks.

	Block 1	Block 2	Block 3
Block 1	0.04796	0.07958	0.06932
Block 2	0.07958	1.00000	0.10000
Block 3	0.06932	0.10000	1.00000

SE provides a partitioning of the network into groups of "similar" actors. For the PPML it provides three separate groups: two small, active groups and one large, relatively inactive group. These groupings provide some insight into the communication patterns between users, and are of keen interest in determining the presence/absence of subgroups of interest in the network. The small groups may represent an active sub-group of experiential knowledge sharers, who actively communicate with all other members of the group. Further investigation of the content of the communications within these two groups is required.

6.2 Visualization

Figure 8 is a capture of the VECoN system. The left pane contains the network visualization, and the right pane contains the spring-embedding controls, the filter controls and the message pane where the contents of the threads or the participating threads are presented.

7 DISCUSSION

The SNA provided some useful insight into the communication patterns within the network, but further analysis is required to fully flush out what the results mean in the scope of the project.

7.1 Isolates

There were 68 isolates recognized in the actor network. These were actors whose messages received no response on the mailing list. A further investigation reveals that some of these actors posted few times to the network; 13 only posted to the PPML once. Looking through the subject lines, there are several messages that are not meaningful pediatric pain communications (such as incorrect subscribe/unsubscribe messages and job or conference announcements). However, there are also meaningful messages which received no response. There does not seem to be a pattern to which messages were ignored, and after filtering the spam messages there were only nine meaningful queries that were left unanswered. This is a positive finding for a mailing list, as it is evidence that knowledge seeking queries are being answered most of the time. Moving forward it is imperative that the PPML continues to be an active community by incorporating new members and responding to their queries.

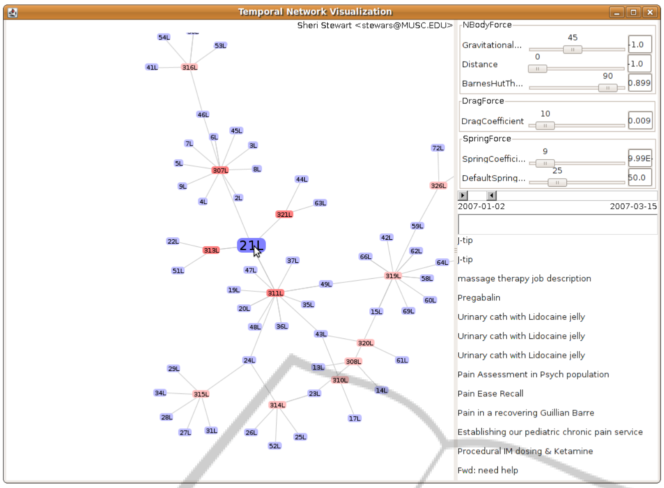


Figure 8: The PPML visualization. Note that the current network is restricted to messages between 2007-01-02 and 2007-03-21.

7.2 Centrality

The centrality analysis has indicated that there is not a central user or set of users that control the communications on the network. For each of the four measurements the highest scoring actors do not have disproportionately high normalized values, indicating that there is not one actor that communicates on every thread, or connects disparate groups of actors. Though the low degree centrality scores could be interpreted as members being inactive, a more reasonable interpretation is that there are many messages on the mailing list that are spawning conversations between different actors. This is a very promising result moving forward.

Detecting content experts or knowledge sharing activities strictly using graph-theoretic principles such as centrality analysis poses problems. If someone contributes to the list by asking many questions, without providing answers, they are recognized as central users, but should not be considered content experts. In order to improve the prediction of content experts it is necessary to extract the underlying semantics of the messages being communicated. Previous work, (Stewart et al., 2010), has worked on extracting semantic representations of the messages, and incorporating this information into the SNA could help differentiate the content experts from the "question askers."

7.3 Structural Equivalence

The structural equivalence analysis broke the model into 3 separate groups: two with high density communications and one with low density communication. Further analysis into the content of the two groups reveals little to identify the two blocks, and there is little to suggest that there are any specific sub-specialities of interest within these two groups. Further investigation should pursue other measures of SE, including incorporating the 2-mode nature of the data.

7.4 Visualization

A beta release of the *VECoN* system has been produced, but it has not yet been tested by the PPML members. Once the mailing list is available as a discussion forum and links between the visualization and the forum are established the *VECoN* system will be rolled out as a Java applet, and at that point research will be conducted on its utility and on future additions. Currently future research is being conducted to improve the spring-embedding algorithm and the blockmodelling algorithm, along with adding more centrality measures and a more intuitive control panel.

8 CONCLUSIONS

Experiential healthcare knowledge is a vital component of the current healthcare system, and developing new methods to facilitate the sharing of this knowledge is vital to sustaining and improving the medical community. Medicine 2.0 technologies provide online tools for facilitating knowledge sharing, establishing virtual communities of clinicians. Understanding the flow of knowledge in these virtual communities is key to developing new systems, and SNA provides the necessary tools for understanding the flow of communication within the network. It has provided a list of potential content experts within the list, it has recognized several active subgroups, and it has partitioned the network into disparate groups of potentially different clinicians.

Further research should be directed at better understanding the members of the community. More actor attributes, such as specialty, location, job description, et cetera, would provide better insight into the structure of the network, and in particular into the structure of the subgroups revealed through blockmodelling. Incorporating the semantic information of the communications themselves would help differentiate between acts of knowledge seeking and knowledge sharing, and improve our overall understanding of the experiential knowledge available in the network.

Though the *VECoN* project is only in its beta stages, preliminary results are promising. The network has been visualized, and SNA tools have been added. Future work will be on adding new SNA tools, improving the visualization, implementing the system online and making the interface more intuitive. The ultimate goal of the *VECoN* system is to provide a novel network exploration tool to help users make new connections within the PPML community, and find new sources of experiential knowledge already available in the network.

ACKNOWLEDGEMENTS

This work is carried out with the aid of a grant from the International Development Research Centre, Ottawa, Canada. The authors would like to acknowledge Dr. Allen Finley for his contributions to the PPML and his ongoing support of this research.

REFERENCES

Barnes, J. and Hut, P. (1986). A hierarchical o(n log n) force calculation algorithm. *Nature*, 324.

Borgatti, S., Everett, M., and Freeman, L. (2002). Ucinet for windows: Software for social network analysis. Harvard, MA: Analytic Technologies.

Borgatti, S. P. and Everett, M. G. (1997). Network analysis of 2-mode data. *Social Networks*, 19:243–269.

Freeman, L. C. (1999). Visualizing social networks. Web. http://www.cmu.edu/joss/content/articles/volume1/Freeman.html.

Fruchterman, T. M. J. and Reingold, E. M. (1991). Graph drawing by force-directed placement. *Software-Practice and Experience*, 21:1129–1164.

Handcock, M. S., Hunter, D. R., Butts, C. T., Goodreau, S. M., and Morris, M. (2003). *statnet: Software Tools for the Statistical Modeling of Network Data*. Seattle, WA. Version 2.1 . Project home page at http://statnetproject.org.

Hanneman, R. A. and Riddle., M. (2005). *Introduction to social network methods.* University of California, Riverside, Riverside, CA.

Heer, J. and boyd, d. (2005). Vizster: Visualizing online social networks. *InfoVis*.

Heer, J., Card, S., and Landay., J. (2005). prefuse: A toolkit for interactive information visualization. In *CHI*.

Kamada, T. and Kawai, S. (1989). An algorithm for drawing general undirected graphs. *Information Processing Letters*, 31:7–15.

Krempel, L. (1999). Visualizing networks with spring embedder: Two-mode and valued data. *American Statistical Association, Proceedings of the Section of Statistical Graphics*, pages 36–45.

Mutton, P. (2004). Inferring and visualizing social networks on internet relay chat. *InfoVis*, pages 35–43.

Nardi, B., Whittaker, S., Isaacs, E., Creech, M., Johnson, J., and Hainsworth, J. (2002). Contactmap: Integrating communication and information through visualizing personal social networks. *Communications of the ACM*, 45(4):89–95.

Stewart, S., Abidi, S., and Finley, A. (2010). Pediatric pain management knowledge linkages: mapping experiential knowledge to explicit knowledge. *Studies in Health Technology and Informatics*, 160:1184–1188.

Wasserman, S. and Faust, K. (1994). *Social Network Analysis: Methods and Applications*. Cambridge University Press.