

# Searching for Sense of Virtual Community in Social Hypertext

Alvin Chin  
Department of Computer Science  
University of Toronto  
Toronto, Ontario, Canada

## **Abstract**

*The role of social hypertext as an indicator of interaction and participation within an associated community is explored. Social network analytic measures are compared with the sense of virtual community that people interacting in several social hypertexts experience, as measured by a well known sense of virtual community questionnaire. The results of these case studies will be interpreted in terms of how well social network analytic measures of virtual community can identify and measure both the different types of community feeling within social hypertexts, and the corresponding amount of sense of virtual community experienced by different participants within the social hypertext.*

## **Alvin Chin**

University of Toronto  
40 St. George Street, Room 8171  
416-978-7581  
achin@cs.toronto.edu  
<http://www.imedia.mie.utoronto.ca/~achin>



PhD-Supervision: Mark Chignell

## **Problem Statement and Research Question**

The problem that I am addressing in my PhD thesis research is how to find, measure and validate relationships between social hypertext structure and corresponding *sense of virtual community*. This sense of virtual community is hypothesized to be somewhat reminiscent of feelings that one would get whilst living in a physical community, and is expected to sometimes occur when interacting in a virtual hypertext environment.

Identifying structural patterns that are indicative of community will help in discovering and tracking social behaviour and social trends. Companies have lots of data about themselves, their customers, their competitors, and their partners, and data comes from various sources such as blogs

and wikis. Methods are needed to make sense of these vast amounts of social data, and to mine connections in the virtual web space for potential customers and competitors.

The central research question in my PhD thesis is how to automatically identify evidence of virtual communities in social hypertext, and how to measure the degree of affiliation of people interacting in virtual environments. I would like to examine to what extent that social hypertexts, based on interactions around blogs, wikis, or photos exhibit properties of virtual community, and how I can validate and measure the extent to which people interacting through those social hypertexts experience a corresponding sense of virtual community.

The main contribution of my dissertation research is a framework and methodology for assessing the amount and type of virtual community based on social hypertext structure. This framework is also used to measure virtual community in a number of canonical examples (case studies) of social hypertext. It is expected that differences in measured sense of virtual community between social hypertexts studied in this thesis will provide initial answers to the following questions:

- Which types of social hypertext media are better for creating and fostering a sense of virtual community?
- What different types of virtual community can be identified based on social hypertext structure?
- What are the properties of the different types of virtual community?
- To what extent can measures of sense of virtual community be inferred from the structural properties of social networks derived from the interactions within those communities?

## **Approach and Methodology**

The first step in my research approach involves creating a theoretical framework and model for sense of virtual community. Properties of this model are then mapped to various social hypertext media through experiments. The approach is based on a *social hypertext model* where social networks of users can be inferred from the interactions that exist within the social hypertext when they respond to other users. The proposed approach for finding sense of virtual community in social hypertext media adapts the concept of *sense of community* characterized by McMillan and Chavis [12]. Social network analysis is used to interpret the structural properties of the social network that can be inferred from the interactions between people in the social hypertext (e.g., in the case of a community of blogs, the interactions may be links between blogs that result from comments made in response to posts). I propose a model for finding sense of virtual community that maps the characteristics of sense of community onto structures that can be inferred from social hypertext networks, using network analytic measures such as *network centrality* [9]. Figure 1 illustrates the methodology for evaluating sense of virtual community in social hypertext.

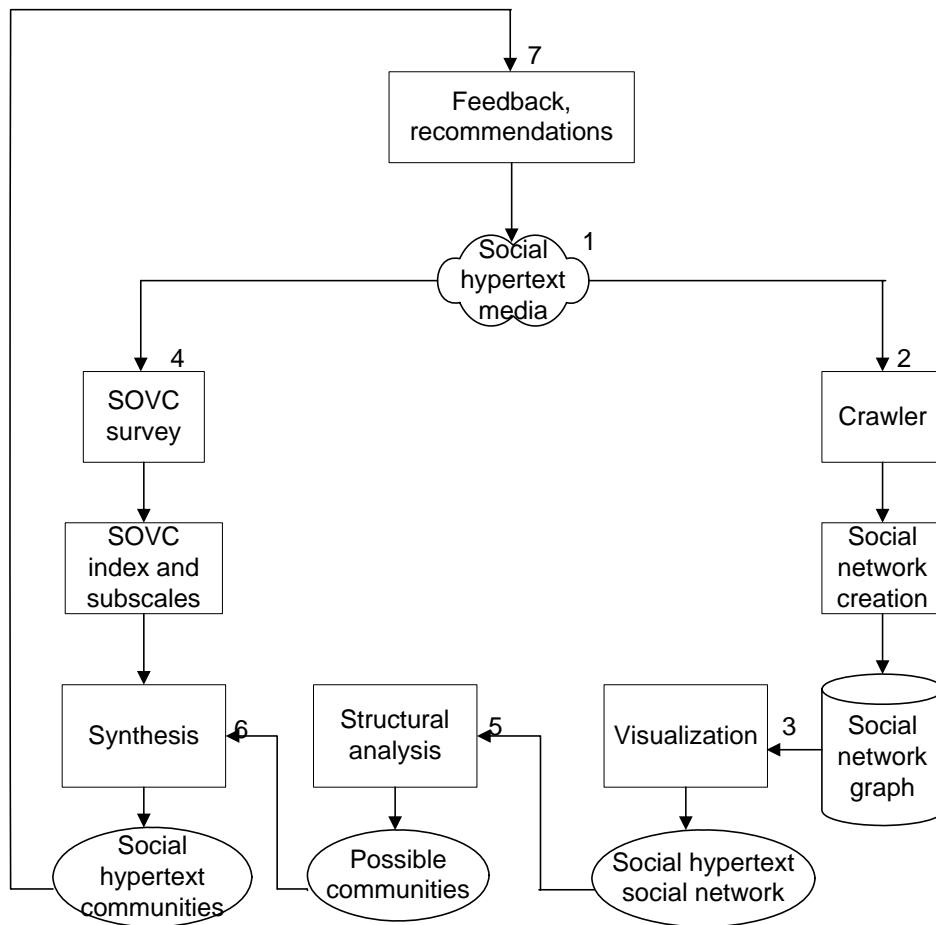


Figure 1. Methodology for identifying and calibrating sense of virtual community in social hypertext.

Step 1 involves finding candidate samples of social hypertext and examining them for possible signs of community. For instance, blogs that have the potential to motivate bloggers to create a community are explored in our earlier work [5]. Once candidate virtual communities have been identified in Step 1, they are crawled in Step 2. The links between people interacting through the community are inferred and recorded as links within a social network that reflects how people interact with each other in the virtual community. In Step 3, key aspects (e.g., centrality) of the social network are measured, and for nodes/people within that network, social network analysis and related methods are used. Visualizations are also constructed to provide a way of viewing the community structure. In step 4, a sense of virtual community survey is administered that asks questions to the users of the social hypertext media about their feelings of sense of virtual community. We adapt the concept of sense of community to create a score for sense of virtual community according to the characteristics of sense of community [2, 12]. In step 5, the structures from social hypertext are mapped to sense of virtual community and network analytic measures of centrality are used to determine if, and to what extent, aspects of virtual community exist within the social hypertext. In step 6, the sense of virtual community scores from the sense of virtual community survey, are used to refine the possible communities into social hypertext communities. Finally, in step 7, the set of identified communities is provided as feedback to the candidate social hypertext media, and used to recommend new links within the social hypertext in order to grow communities within the social hypertext. It is important to note that use of the sense of virtual community questionnaire referred to in step 4 of Figure 1, is intended for initial calibration of the

system for measuring sense of virtual community based on structural properties of social networks. In later practice, it is expected that structure-based predictive measures of sense of virtual community may be sufficient (without the use of questionnaires). In this approach, the automated structure-based approach might provide useful feedback to revise and stimulate interactions within social hypertexts so that they become stronger virtual communities.

In this dissertation research, communities that are discovered from structural analysis of social hypertext, are validated using a sense of virtual community survey that scores based on the subscale of each characteristic [4]. Automatically measured sense of virtual communities based on inferred social networks are compared to measures of sense of virtual community within social hypertexts collected using the sense of virtual community survey, in order to iteratively improve on the algorithm for identifying virtual community.

The sense of virtual community questionnaire is used because analyzing social hypertext requires not just computational analysis of its links, but also understanding of the social behaviour that underlies the creation of those links. The sense of community model from McMillan and Chavis [12] as applied to virtual community captures this.

Other alternatives to our approach that may supplement the social network analytic approach include using clustering algorithms to find groups of related nodes or clusters within the social hypertext such as blogs [1]. However, there are many different types of clustering algorithms [10], therefore research will be needed to discover which types of clustering might be useful in assessing virtual community.

## **Related Work**

The Web is a network of hypertext which forms an “association of thoughts, in accordance with some intricate web of trails”, according to Bush [3]. Commentators on early hypertext systems predating the Web observed that the links between pages in hypertext can form relationships to people, places and things in the real world, and can establish online presence and social identity. Since web pages are created by people, and the content is provided for others to search and navigate, the relationships become links within a social network. Erickson was the first to call this “*social hypertext*” [8] because people start to portray themselves to others on the Web, and they start engaging in social relationships around shared content and interests.

Computer networks can be represented and analyzed as social networks (Wellman [15]). When web pages involve people, social networks start to emerge because each web page is represented as a node in the social network. Numerous researchers including Watts [14] have studied the social networking structure of computer networks which has led to a number of structure-based innovations, including prominent searching and ranking algorithms used by Google and others.

Finding community in virtual (as against physical) environments such as e-mail and the Web is difficult because there is no physical location and no physical human interaction to observe. Some researchers question whether communities can be found in virtual environments, since there is a lack of identification with place [7]. Community in virtual environments does not relate to specific places, but can be identified by common ties and social interaction. There is a wealth of literature demonstrating that such virtual community exists in virtual environments such as [16] and that virtual community can support ties to a large number of people, where those ties would otherwise

not exist. According to Rheingold [13], virtual communities are "social aggregations that emerge from the Net when enough people carry on those public discussions long enough, with sufficient human feeling, to form webs of personal relationships in cyberspace".

Networks of hyperlinked pages can also be analyzed using clustering algorithms such as those reviewed in [10]. The seminal work of John Kleinberg [11] paved the way for searching algorithms based on clustering on the Web. Kleinberg was one of the first to use link analysis combined with an iterative algorithm to compute these clusters, which he called *hubs* and *authorities*. Link analysis is the method for analyzing a network based on the links between the nodes, and is commonly used to identify topics within the clusters of pages on the Web that are strongly linked to each other. Kleinberg created an algorithm called HITS (Hypertext Induced Topic Search) where authorities are collections of pages that are linked to by many other pages, while hubs are pages that point to authorities. The strongest authorities can be identified as communities since they can be considered experts on a topic.

While researchers are recognizing the role of hypertext in building social networks, the correlation between users' feelings and structured properties of social hypertext has yet to be addressed. Research in virtual community started by Rheingold [13] and Wellman [15, 16] do not exploit the structure and relationships of links that are recorded in the accompanying social hypertext. As a result, this leads to multiple interpretations of virtual community. Even though clustering and network analysis are increasingly being used as a method for identifying community, there exists no universal consensus as to which measures may be most closely related to different aspects of community.

In order to find and measure a sense of virtual community in social hypertext environments, there needs to be an understanding of the construct of sense of virtual community and how much of that construct can be identified and measured. We take a multidisciplinary approach to address this deficiency, using a combination of behavioural measures, hypertext crawling, and social network analysis. In particular, we adapt the concept of sense of community introduced by McMillan and Chavis [12] as the underlying definition for community, because community is a psychological and social entity, and has been used by many researchers for finding community [2]. Previous work has looked at identifying structures within social hypertext using clustering techniques [1], but without validating whether those structures were virtual communities. The research in this dissertation will extend earlier research results by creating a social hypertext model that maps the features of sense of virtual community to the structural patterns within the social hypertext network and to quantitative measures from structural analysis.

## **Preliminary Results**

An enhanced understanding of virtual community can be used to identify and supplement missing human and social context in pervasive computing. We anticipate that trends towards increased use of pervasive computing will increase the frequencies of interactions within social hypertexts, thereby increasing sense of virtual community. In this view, use of pervasive computing will be an accelerator for the formation and growth of virtual communities.

Our model has been applied to a blog [6] which was created to supplement a Canadian independent music web site for tracking the beginnings of virtual community forming around a new blog. The network of reciprocally linked blogs was found to form a set of connected stars which were

hypothesized to be typical of blogs in an early stage of development. We conducted a simple analysis involving six blogs which were authored by readers of the independent music blog that was the centre of the community under consideration. Three of the blogs showed evidence of being part of the community, in terms of the evidence of community formation visible in the network visualization, and in terms of both the social network analytic centrality measures used, and the strength of community scale and its subscales.

We are carrying out an information validation of the results by comparing the quantitative measures of sense of virtual community using network centrality with the visual structural patterns and frequency of links. By asking participants for their feedback to our identification of virtual communities, we can iteratively improve on and validate our social hypertext model and algorithm.

## **Conclusions and Future Steps**

Current work involves applying the social hypertext model to a richer and larger dataset of social hypertext media, like blogs, newsgroups, and tagging systems, and conducting a larger sample of participants for measuring sense of virtual community. Future work will involve developing a rigorous model and quantitative framework for automatically assessing and measuring sense of virtual community in large collections of social hypertext, further developing a measure for community that uses network centrality as a parameter, and validating the sense of community model by conducting post-interviews with users. One of the challenges that we are currently addressing is obtaining a good enough dataset that shows the potential for finding sense of virtual community and exhibits characteristics of a sense of virtual community. Another challenge is to persuade a significant number of the users to complete the sense of virtual community survey.

We need additional and external expertise on datasets and case studies that we could use for testing our social hypertext model and conducting the surveys. We would like to cooperate with others who may have other techniques and suggestions of algorithms for automatically discovering communities from the network graphs, and for conducting behavioural analysis. We would appreciate guidance and advice on what would be a reasonable and practical definition and measure for community in the context of social hypertext media, as well as particular techniques for validating our social hypertext model analytically.

## **References**

- [1] E. Adar et al. Implicit structure and the dynamics of blogspace. Workshop on the Weblogging Ecosystem, 13th International World Wide Web Conference, 2004.
- [2] A. Blanchard and M. Markus. The experienced sense of a virtual community: Characteristics and processes. *The DATA BASE for Advances in Information Systems*, 35(1), 2004.
- [3] V. Bush. As we may think. *The Atlantic Monthly*, 176(1):101–108, 1945.
- [4] D. Chavis. Sense of community index. <http://www.capablecommunity.com/pubs/SCIndex.PDF>.
- [5] A. Chin and M. Chignell. Finding Evidence of Community from Blogging Co-citations: A Social Network Analytic Approach. In *Proceedings of 3rd IADIS International Conference Web Based Communities 2006*. San Sebastian, Spain, 191-200, 2006.

- [6] A. Chin and M. Chignell. "A Social Hypertext Model for Finding Community in Blogs". In Proceedings of the 17th International ACM Conference on Hypertext and Hypermedia: Tools for Supporting Social Structures, Odense, Denmark, 11-22, 2006.
- [7] R. B. Driskell and L. Lyon. Are virtual communities true communities? Examining the environments and elements of community. *City and Community*, 1(4):373–390, 2002.
- [8] T. Erickson. The world-wide-web as social hypertext. *Communications of the ACM*, 39(1):15–17, 1996.
- [9] L. C. Freeman. Centrality in Social Networks: Conceptual Clarification. *Social Networks* 1: 1978/79, 215-239.
- [10] J. Hartigan. *Clustering Algorithms*. John Wiley and Sons, New York, NY, 1975.
- [11] J. M. Kleinberg. Authoritative sources in a hyperlinked environment. *Journal of the ACM*, 46(5):604–632, 1999.
- [12] D. W. McMillan and D. M. Chavis. Sense of community: A definition and theory. *Journal of Community Psychology*, 14(1) 6-23. 1986.
- [13] H. Rheingold. *The virtual community: Homesteading on the electronic frontier*. Addison-Wesley, Toronto, ON, Canada, 1993.
- [14] D.J. Watts and S.H. Strogatz. Collective dynamics of "small-world" networks. *Nature*, 393:440–442, 1998.
- [15] B. Wellman. Computer Networks As Social Networks. *Science*, 293(5537):2031–2034, 2001.
- [16] B. Wellman and M. Guilia. Net surfers dont ride alone: Virtual communities as communities. In B. Wellman (Ed.), *Networks in the global village: Life in contemporary communities*, 1999.