A Unifying Framework for Building Social Computing Applications

Renato A.C. Capuruço and Luiz F. Capretz

University of Western Ontario, Department of Electrical and Computer Engineering
London, Ontario, Canada, N6A 5B9
`{r.capu,lfcapretz}@uwo.ca`

Abstract. There have been a number of frameworks and models developed to support different aspects of social computing. Some were developed to deal with online interaction through the application of computer-mediated communications tools, whereas others such as social network analyses and reputation systems were more specific in their focus. While these methodologies are inter-related, current social computing research has dealt with them as separate aspects. This paper presents a comprehensive framework for social computing that aims at integrating all these three aspects into a unified model so that discovery and exploration of community members are not only made possible, but also optimized. The novelty of the proposed approach stems from: (1) integration of social interaction, social network analysis and social reputation domains; (2) incorporating many flexible and practical features such as individual- and group-level perceptions of trust in social relations for different social contexts; and (3) using path-related algorithms for selection and discovery of target community members. A description of the proposed model and its implementation are presented. This research is expected to assist online community members to make decisions that facilitate the discovery of people and their connections while promoting increased awareness of community structure and information exposure.

Keywords: Social Computing, Social Informatics, Reputation Systems, Online Interaction, Social Networks, Web 2.0.

1 Introduction

In recent years, social computing has received considerable attention in North America and worldwide. The proliferation of online social networking services, in which millions of members publicly articulate mutual "friendship" relations has given rise to many forms of online sociality. These tendencies to form online social groups or live in an online community have also powered the rise of social computing upon fundamentals such as computer-mediated communication tools, reputation systems and social network analyses. While online social interactions at these three levels are inter-related, social computing research has dealt with them as separate aspects; thus, may lead to neither being articulated in cooperation one to another to foster more compelling nor more effective social interactions.

Based on the theoretical constructs of sociology and mathematical foundations of graph theory, social network analysis offers a unique methodology for visualizing and investigating social structures and relations [1]. Social network analysis also produces an alternate view of complex sets of relationships between members of social systems at all scales, where the relationships and ties of an individual with others are more important than their personal attributes alone. From these analyses and mathematical models streams that are used to understand and analyze the social network data, several generalizations about the features of personal networks have emerged and a summary of the most relevant ones can be found in [2, 3, 4, 5, 6, 7 and 8]. Granovetter [2, 3] has revealed, emphasized and shown the importance of different kinds of ties (e.g., weak, strong) have on individual and communities accessing their network resources for a purpose, such as finding jobs. Blau [4] and Burt [5] have analyzed the importance of social exchanges in interpersonal relationships of multi-group affiliations to reveal the ultimate effects of ethnic, socioeconomic, and other aspects of population structures to achieve societal power, and large-scale domination (economic or political) of groups without personal contact. Moreno [6], for example, introduced the basis of ever-evolving quantitative methods for measuring social relationships, but it was Freeman [7] who showed that although over the years a great many measures of centrality – an indicator to identify a person's position in the network, who is capable of lower or higher network influence [8] – have been proposed, those are often only vaguely related to the intuitive ideas that support the index, and many are so complex that it is difficult or impossible to discover what, if anything, they are measuring. Despite of the increasing trends in developing theoretical studies on social networks [9] and latest developments in applied computer applications for social network analysis [10], existing research still has some drawbacks related to modeling complexity of relationship phenomena that enables an individual to achieve societal gains, which the present research is attempting to overcome.

Bringing physical communities to the virtual world is a complex issue. Several design principles for successfully establishing them in the later environment have been underlined. For example, Kollock [11] suggested that a flourishing online community has to (i) create the condition that two individuals will meet again in the future, (ii) create the condition that individuals must be able to identify each other, and (iii) create the condition that individuals must have information about how the other person has behaved in the past. In addition, Godwin [12] stresses the importance of promoting continuity in online groups and that online communities should provide institutional memory – durable records of the events and history of the group, and Rourke et al. [13] anticipated that the ability for community members to leave comments in an interactive format is an important element to support the cognitive and affective objectives of social interaction. Computer-mediated communication techniques such as blogs and chat rooms can help communities meet those challenges. They are found to be the most popular and accepted interaction tools to constituting and maintaining an online social presence [14], thus becoming a key ingredient in constructing online communities for the purpose of this research.

Kollock [15] has also stated that reputation and trust are the bedrock of community ongoing interaction and cooperation, and are a vital source of social information and control. In this regard, considerable amount of research that has focused on the development of trust and reputation models as the most accepted and popular methods to

capture trustworthy in the online interaction process can be found in [16, 17, 18, 19, 20, 21, 22, 23 and 24]. There are two main streams of reputation systems: e-commerce and peer-to-peer applications. In Gupta et al. [16], peer-to-peer applications including KaZaA[1] are examined and a trust model is proposed where different parameters such as the average query-response message size, the ration of Mbytes uploaded, and the amount of content-shared are used for computing the reputation score associated with peers. On the other hand, Dellarocas et al. [21] is an example of research on reputation systems that are largely used only for online trading communities, such as e-Bay[2]. The reliability of participants in such environments is measured by calculating a score associated with one or more of a user's participation level (e.g., number of successful transactions), availability of physical identities (e.g. valid email), and feedback about interactions with each other.

However, incorporating reputation information into social networks formed from social interactions other than the ones derived from e-commerce- and peer-to-peer-based communities complicates and renders traditional methodologies as insufficient to deal with the distinct formulation involved. A new methodology is needed, which the present research is also attempting to overcome.

While existing social computing methodologies and models have dealt specifically with single aspects of social interaction, social reputation, and social networks, they still have some drawbacks that may hinder the practical discovery and exploration of community members in terms of social spaces, casual interactions, and meaningful exchanges, particularly when considering all three domains combined.

This paper presents a comprehensive and transparent model to assist online community members to make decisions that facilitate the discovery of people and their connections while promoting increased awareness of community structure and information exposure. This is achieved by integrating the social interaction, social network analysis and social reputation domains; by incorporating many flexible and practical features (e.g., social contexts, social priorities, social relations, and social intensities); and by incorporating path-related algorithms for selection and discovery of target members. The framework developments and its implementation on a prototype application are outlined. The output of the model is a set of search and guidance strategies based on individual and group perceptions of trust in social relations for different social contexts for the discovery and exploration of community members throughout the underlined social network.

2   Components of a Unified Social Computing Framework

The main components of the unifying social computing framework (SCF) that incorporates the three above-mentioned domains are as follows:

- Detailed SFC models (online social interaction tools, social network, and reputation-dependent perceptions of qualities or attributes)
- SFC Constraints (social context, social relation, social reputation, user defined constraints such as social relation intensity, priority etc.)

---

[1] www.kazaa.com

[2] www.ebay.com

- SFC Decision Support Module (user interface, community database, social relation assessment, reputable search, visualization)

2.1 Framework Models

At the core of a successful SCF are proper models for promoting interaction and capturing the relationship patterns of the individuals in the online community so that several interaction patterns can be estimated and the benefits of knowing those returned to the community itself.

Interaction Model

The SCF requires an online interaction model that supports online human interaction and information flow so that communities are formed for ongoing collaboration and exchange of information and knowledge among their members. In this regard, people form online communities by using a combination of one-to-one (e.g., instant messages, e-mails, chat rooms), one-to-many (e.g., web pages and blogging), and many-to-many (e.g., wikis) communication modes [25]. This model is the entry point in which personal networks are formed.

Social Network Model

The social network model represents the logical structure that embodies the patterns of the relationships between social actors (i.e., members or individuals with a particular role of an online community) at all scales and the possible statements that can drawn from those by using social network analyses-based techniques.

In this research, a combination of Laumann et al.'s [26] three generic approaches to decide on the set(s) of objects that lie within a social network and Scott's [27] definition of the principal types of data to be considered to fulfil those approaches was employed. As such, this model uses two types of data as the model's building blocks: attribute- and relational-based data. According to [27], Attribute data relates to the attitudes, opinions and behaviours of objects (i.e., actors) combined with their basic characteristics to define formal membership criteria. These data sets are regarded as the properties, qualities or characteristics that belong to them as individuals or groups. Relational data, on the other hand, are the contacts, ties, and connections, the group attachments and meetings, which relate one actor to another and so cannot be reduced to the properties of the actors, but of systems of actors; these relations connect pairs of actors into the larger relational system.

The two basic types of data are translated into the community members' profile and their connections features. While a set of socio-demographic characteristics such as age, gender, etc. is the most natural dataset candidate to be organized as attribute data, an assortment of opinions representing expressions of the experience when dealing with particular actors can be also structured as attribute data. The collection of relations connecting pairs of individuals such as "friend of whom", "has studied with", "has worked with", etc. emerges as specific community-generated content that can be mapped as relational data.

Social Reputation Model

In this paper, the reputation model is designed to take into account both individual and group perceptions for the person in which others are linked. Sabater et al. [28]

suggested that reputation is not just based on facts, but also based on other's beliefs about the subject of the reputation calculation.

Therefore, the perception of trust in this model can be divided into four parts: (i) the category the reputation information belongs to, (ii) the amount of reputation (i.e., rating) assign to a particular category, (iii) the feedback type used to collect those judgments, and finally (iv) the relative importance (i.e., weight) among the categories.

These perceptions of trust are aggregated into a numerical value, which synthesizes the social impressions of interaction quality and trust that not only a person has about another, but also perceptions the community as a whole has about an individual.

At the group level, candidate rating parameters such as *frequency of online partici-pation*, *willingness to share private information*, and *being recommended by other members* allows the calculation of a Social Reputation score (Eq. 1) for each member of the community. The score can be automatically determined as a weighted sum of the reputation ratings of each of the categories, considering the respective category's intensity (weights), and dividing them by the sum of the weights. In the scope of this research, this score represents the total informed judgment on the trustworthiness of a member by the community he or she belongs to.

$$SRS_{group} = \frac{\sum \left(Reputation\ rating \times Intensity \right)}{\sum Intensities} \qquad (1)$$

At the individual level, candidate rating parameters such as *cognitive*, *interactive*, and *affective* types of relations allows the calculation of a Social Relation score (Eq. 2) for the connection existing between each pair of individuals of the community. The score can be automatically determined as a weighted sum of the relation ratings of each of the relation types, considering the respective relation intensity (weights), and dividing them by the sum of the weights. This score represents the total intuitive judgment on the interaction of a pair of community members.

$$SRS_{individual} = \frac{\sum \left(Relation\ rating \times Intensity \right)}{\sum Intensities} \qquad (2)$$

The idea behind the score weights is to allow the community members to reflect on their unique intuitive knowledge about what category matters the most, which can vary from one member's experience to another.

2.2 Framework Constraints

Several practical constraints should be taken into consideration for implementing social-aware applications. These constraints can be categorized as follows: *social context*, a specific and common relation environment among pair of people in which social interactions happen; *social relations*, the different types of interactions among pair of people; *social reputation*, the measure of judgments and perceptions about the character, stability, reliability, behaviour, etc. of people who interact in a given community; *relation intensity*, the measure of the relative strength, importance or "bond energy" among interactions; *relation priority*, the measure of importance of the

relationship among pair of people; *user*, the ability of a community member to enforce his or her decision on the decided one; *privacy and security*, the governing policies that allow disclosure of personal information and access privileges to those. These are important aspects to be considered in the design of a general social computing system.

2.3   Framework Decision Support

The SCF decision support integrates both the SCF models and the SCF constraints to arrive at the social-aware application. The SCF decision support component comprises of a reputable search optimization model linked to the portfolio (database) of communities and to a social relation assessment model that applies all the SCF models to all components.

Community Database

The underlined physical structure of the community network that supports the social network model is a familiar database schema based on node-link representation, where nodes represent members of the community and links denote the articulated "social relationships" (e.g., interaction, ties) between them. Each node and link has attributes associated that allows users to calculate and store reputation information. Each link is associated with a particular social context, and a pair of nodes may have one or more links, thus representing different social contexts of interactions.

Based on the both profile and connection attributes, a relational database management system was designed and two main tables – nodes and links, respectively – were implemented to store, in real-time, the network objects and the associations between objects. This network model is the working data repository that becomes available for further processing by the decision support engine.

Social Relation Assessment

A social relation rating system had to be developed to perform the condition assessment of the social reputations and social relations in the network. The condition rating used in this methodology varies according to the reputation category and relation type. Generally, the rating uses a general scale from 1 to 5 for the reputation elements. This scale assumes that social reputations and relations are valued as the worst and the best, respectively. Condition ratings are used to describe the existing condition of trust and opinions among individuals in the social network. It is considered as the most important phase on which subsequent decisions are based.

The Social score calculation mechanism works as a function of the type of user feedback: when the online community is capable of storing complete and accurate information about the transactions they mediate (e.g., number of logins per member) to execute the calculation, this is aggregated automatically by batch routines, without user's direct participation. On the other hand, when explicit input from the rest of the community members is needed to express comments about interactions with each other, then the calculation is performed manually, upon user's request for participation (e.g., online voting).

Reputable Searchability using Path-related Algorithms

Having defined the present condition of a social network with the online interaction model and reputation model, the proposed SCF uses a Path-based optimization model

[29] to determine optimum priority list of members and their social relations conditions. *Reputable Searchability* is a newly coined term to define the class of social computing search engines that are capable of show a target member based on a desired level of individual- and/or group-based reputation. *Reputable Searchability* is very important to consider because a member of a community is not defined by its ethnographic attributes only; rather, he/she is characterized by a combination of those with his or her social ties [30]. The reputable search engine also optimally "guides" the searcher to its desired destination functioning as a means or medium for showing all linkages between two or more people. As such, the procedure searches for the path with lowest cost between a community member and every other member with respect to user defined constraints.

In order to develop a sound reputable search mechanism, a Social Relation Index as a combination of group- and individual-based perceptions was constructed to account for the social distance between a pair of members, as per Equation (3) below:

$$SRI = \frac{\left(SRS_{individual} \times Weight_i\right) + \left(SRS_{group} \times Weight_g\right)}{\sum Weights_{i,g}} \qquad (3)$$

The Social Relation Index (SRI) is calculated as the weighted sum of the individual (i.e., Social Relation Score) and group (i.e., Social Reputation Score) scores, considering the relative importance between them (i.e., weight), and dividing the result by the sum of the weights. This index represents the total "social distance" between a pair of members. The "social distance" is the basis for the cost structure that had to be implemented, so that the path-related algorithm can function accordingly.

Implementing a graph search algorithm in the framework involves four main steps: (1) eliciting community members and their connections for a given social context; (2) setting the source and destination nodes; (3) deciding on the evaluation criteria, higher or lower SRIs; and (4) applying the relaxation principle to generate short paths. The cost structure is a function of the Social Relation Index and because lower scores mean worst case scenarios (e.g., lower affectivity, interactivity, etc.), the procedure had to be adapted to find the path with the highest cost as well (i.e., higher SRI), more specifically meaning paths with highest quality in terms of social relation (e.g., higher affectivity, interactivity, etc.).

After defining the cost structure, the constraints considered in the algorithm are:

1. Choose a Social Context – This will filter out members and/or connections;
2. Choose one or more Social Relations, assigning corresponding weights – This will affect the calculation of the Social Score at the individual level;
3. Choose "Group" or "Individual" reputation – This will affect not only the calculation of the Social Score at the group level, but it will also affect the calculation of the whole Social Index by including/excluding either or both levels;
4. Specify whether to use lower or higher scores – This will affect the selection of target members by the algorithm.

To evaluate a possible solution (list of members), the reputable search algorithm identifies, analyses and builds the cost structure by using the desired constraints and

social context for a particular population (community). Once the target population of that community has been created, the social distance is calculated for each social interaction for all of their members. Then, beginning from the source node (member), paths from one node to another whose total cost is the least among all such paths is calculated until the target node is reached.

Visualization
A visualization model includes representation and presentation features suggested by Carpendale et al. [31]. The model also should support a range of basic exploratory search features by such methods such as panning, scrolling, zooming, etc., providing visualization of the rich profile and connection data characteristics as of traditional "Sociograms" [32].

3   Discussions

The framework model presented in this paper has been demonstrated to work effectively on the example application. Further experimentation was conducted on different combinations of personal networks with different properties, and the model proved to consistently produce expected results. The exploration and discovery of members of a community by using the proposed framework is a powerful feature that brings along the necessity to discuss its implications for the knowledge society as a whole.

   In today's electronic age, raw data has become a valuable commodity and the protection of personal information has become increasingly important to our sense of privacy. New technologies such as the framework being proposed will certainly create challenges for the protection and use of personal information. Two core approaches could be used to help address this issue.

   First, the model should be designed with privacy tools necessary to control how and with whom personal information is shared from one individual to another. This would give individuals the sense of safety of their information, making the improper collection or misuse of information more difficult to achieve.

   Second, while this technology is not required for an invasion of privacy, the ability of techniques to amplify, routinize and sublimate surveillance to collect and use the health of personal data with the proposed model raises some even greater privacy concerns. For example, practices including the monitoring of telephone calls and computer use could be employed to collect non-disclosed personal data, which could be plugged into the framework to extract commercial and legal value from people's interaction in rich, powerful, and flexible ways.

   In many respects, private sectors, government and citizens groups may have to work together towards effective legal safeguards and proper communication methods to the practical aspects of such technologies and techniques. This is an essential ingredient to promote the responsible development of such methodologies, while protecting information privacy and rights.

4   Conclusions

In this paper, literature related to online interaction tools, social network analyses, and reputation systems has been reviewed and a model is presented to integrate these three aspects into a unified social computing application framework.

The proposed framework incorporates a reputable search engine based on path-related algorithms to calculate the social relations conditions for members of an online community and optimally generate a list of members between any target and destination persons. The developed model is flexible and allows for several customizations for more effective searches. In addition to its expandable data structure, some of the flexible features of the proposed framework that make it an efficient model for building social applications include:

- Combination of three research venues (online social interaction, social network, and reputation systems) into a single methodology;
- Reputable Searchability process with optimization process that respects desirable social distance;
- Incorporate Social Relation Index as indicator to assess the social relation condition of the network;
- Consider two levels of reputation: group and individual;
- Consider variable types of relationships;
- Consider variable categories of social context, one at a time; and,

References

1. Wasserman, S., Faust, K.: Social Network Analysis: Methods and Applications. Cambridge University Press, New York (1994)
2. Granovetter, M.: The Strength of Weak Ties. American Journal of Sociology 78, 1360–1380 (1973)
3. Granovetter, M.: The Strength of Weak Ties: A Network Theory Revisited. Journal of Sociological Theory 1, 201–233 (1983)
4. Blau, P.: Macro Structural Concepts. In: Structural Contexts of Opportunities. University of Chicago Press, Chicago (1994)
5. Burt, S., Ronald: Structural Holes. Univ. Chicago Press, Chicago (1992)
6. Moreno, J.L.: Sociometry: Experimental Method and the Science of Society – An Approach to a New Political Orientation. Beacon House, Beacon (1951)
7. Freeman, L.C.: Centrality in Social Networks Conceptual Clarification. Journal of Social Networks 1, 215–239 (1978/1979)
8. Bavelas, A.: A mathematical model for group structures. Human Organization 7, 16–30 (1948)
9. Freeman, L.C.: The Development of Social Network Analysis: A study in the Sociology of Science. Empirical Press, Vancouver (2004)
10. Huisman, M., Van Duijn, M.A.J.: Software for Social Network Analysis. In: Carrington, P.J., Scott, J., Wasserman, S. (eds.) Models and Methods in Social Network Analysis, pp. 270–316. Cambridge University Press, New York (2005)
11. Kollock, P.: Design Principles for Online Communities. In: The internet and Society. Harvard Conference Proceedings. O'Reilly & Associates, Cambridge (1997)

12. Goodwin, M.: Nine Principles for Making Virtual Communities Work. Wired 2.06, 72–73 (1994)
13. Rourke, L., Anderson, T., Garrison, D.R., Archer, W.: Assessing Social Presence In Asynchronous Text-based Computer Conferencing. Journal of Distance Education/Revue de l'enseignement à distance (2001)
14. Boyd, D.M., Ellison, N.B.: Social network sites: Definition, history, and scholarship. Journal of Computer-Mediated Communication 13(1), article 11 (2007)
15. Kollock, P., Smith, M.: Managing the Virtual Commons: Cooperation and Conflict in Computer Communities. In: Herring, S. (ed.) Computer-Mediated Communication: Linguistic, Social, and Cross-Cultural Perspectives, pp. 109–128. John Benjamins, Amsterdam (1996)
16. Gupta, M., Judge, P., Ammar, M.: A reputation system for peer-to-peer networks. In: Proc. of the ACM 13th International Workshop on Network and Operating Systems Support for Digital Audio and Video, Monterey, California, USA (2003)
17. Damiani, E., De Capitani di Vimercati, S., Paraboschi, S., Pesenti, M., Samarati, P., Zara, S.: Fuzzy logic techniques for reputation management in anonymous peer-to-peer systems. In: Proc. of the Third International Conference in Fuzzy Logic and Technology, Zittau, Germany (2003)
18. Damiani, E., De Capitani di Vimercati, S., Paraboschi, S., Samarati, P., Violante, F.: A reputation-based approach for choosing reliable resources in peer-to-peer networks. In: Proc. of the 9th ACM Conference on Computer and Communications Security, Washington, DC, USA (2002)
19. Dellarocas, C.: Immunizing online reputation reporting systems against unfair ratings and discriminatory behavior. In: Proc. of the 2nd ACM Conference on Electronic Commerce, Minneapolis, MN, USA (2000)
20. Dingledine, R., Freedman, M.J., Hopwood, D., Molnar, D.: A reputation system to increase MIX-net reliability. In: Moskowitz, I.S. (ed.) IH 2001. LNCS, vol. 2137, p. 126+. Springer, Heidelberg (2001)
21. Dellarocas, C., Resnick, P.: Online Reputation Mechanisms: A Roadmap for Future Research. In: Summary report of the First Interdisciplinary Symposium on Online Reputation Mechanisms, April 26-27 (2003)
22. Rahman, A., Hailes, S.: Supporting trust in virtual communities. In: Proc. of the IEEE Hawaii International Conference on System Sciences, Maui, Hawaii (2000)
23. Yu, B., Singh, M.: A social mechanism for reputation management in electronic communities. In: Proc. of the 4th International Workshop on Cooperative Information Agents, Boston, USA (2000)
24. Zacharia, G., Moukas, A., Maes, P.: Collaborative reputation mechanisms in electronic marketplaces. In: Proc. of the 32nd Hawaii International Conference on System Sciences, Maui, Hawaii (1999)
25. Shirky, C.: A Group is Its Own Worst Enemy. In: 2nd Annual O'Reilly Emerging Technology Conference, Santa Clara, California, April 20-25 (2003), http://www.shirky.com/writings/group_enemy.html
26. Laumann, E.O., Marsden, P.V., Prensky, D.: The Boundary Specification Problem in Network Analysis. In: Freeman, L.C., White, D.R., Ronney, A.K. (eds.) Research Method in Social Network Analysis, pp. 61–87. George Mason University Press, VA (1989)
27. Scott, J.: Social Network Analysis: A Handbook, ch. 1-2, p. 2. Sage, London (1991)
28. Sabater, J., Sierra, C.: Review on computational trust and reputation models. Artificial Intelligence Review 24(1), 33–60 (2005)

29. Dijkstra, E.W.: A note on two problems in connexion with graphs. In: Numerische Mathematik, vol. 1, pp. 269–271 (1959)
30. Simmel, G.: The Number of Members as Determining the Sociological Form of the Group. American Journal of Sociology 8(1), 1–46; (2):158–196 (1902)
31. Carpendale, M.S.T., Montagnese, C.: A Framework for Unifying Presentation. In: Proceedings of ACM User Interface Software and Technology UIST 2001, pp. 82–92 (November 2001)
32. Freeman, L.C.: Visualizing Social Networks. Journal of Social Structure (2000)