HAL open science

Modeling Groups in Social Networks

Vincent Levorato

▶ To cite this version:

Vincent Levorato. Modeling Groups in Social Networks. (ECMS), Jun 2011, Poland. pp.129-134. hal-00605395

HAL Id: hal-00605395
https://hal.science/hal-00605395

Submitted on 1 Jul 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

MODELING GROUPS IN SOCIAL NETWORKS

Vincent Levorato
LIFO - Université d'Orléans
Rue Léonard de Vinci
B.P. 6759, 45067 ORLEANS Cedex 2, France
Email: vincent.levorato@univ-orleans.fr

KEYWORDS

complex systems ; modeling ; simulation ; social networks ; group ; pretopology

ABSTRACT

Studies in social network analysis consider the behavior of a group to be equivalent to the "sum" of individual behaviors. However, a group has to be considered as a complex system where emerging phenomena can appear. In this paper, a formalism is proposed to resolve this problematic by modeling groups in social networks using pretopology as a generalization of the graph theory. We also give an example of simulation using our formalism.

Introduction

Network modeling is used in several research domains (physics, biology, computer sciences, ...) and is mainly based on graph theory. Especially in social networks modeling, graphs are used to describe the links representing relationships or flows between entities. The studies consider in most cases individuals as single elements, a group being formed by several persons interacting with each other. Most work on modeling social networks consider a group as a combination of persons (Everett and Borgatti, 2005), not as a whole entity. As social networks are complex networks (Newman et al., 2006), emergence of phenomena can occur, and the behavior of a group of persons can be different from the "sum" of each behavior of the persons composing the group. Thus, in our opinion, graph theory is inadequate to model all complex interactions occurring in a social network. As some group modeling approaches based on graphs have been proposed before (Everett and Borgatti, 1999, 2005), and more recently (Saha and Getoor, 2008), to our knowledge, there is *no general framework for modeling a group of entities in a network*. From this fact, we propose a mathematical framework for group modeling in social networks from a topological point of view.
This paper is structured in two parts: first we introduce pretopology formalism, and we give some definitions and illustrations for group modeling in a social network. Then we show a simple simulation example of a model using our approach. The reader of this paper is supposed to be familiar with graph theory and some concepts from mathematical topology.

Modeling groups in social networks

Before giving theoretical definitions, we first illustrate the problem of a realistic group modeling by giving a simple example: we would like to represent interactions between four persons. A small social network with a simple symmetrical friendship relation shall help in illustrating our approach (Fig. 1).

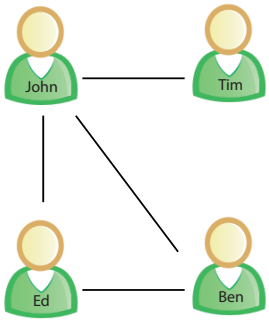


Figure 1: Example of a small social network with friendship relations.

Let's focus on John and Ed. John has three friends: Tim, Ben and Ed, while Ed has two friends: Ben and John. As we notice, it is very easy to know the friends of a person are. But now, we want to know whose people are the friends of the group $\{John, Ed\}$. In this case, having a look at the non-oriented graph induced by the network, we see Tim in relation with John, and Ben in relation with John and Ed. As we can reach Tim and Ben from John and Ed, we could say that Tim and Ben are friends with John and Ed. However, in reality, Ed and Tim are not friends. We need to set a constraint, saying for instance *the friends of the group* $\{John, Ed\}$ *are friends with each person of the group*. So, we finally have only Ben as a friend of the group $\{John, Ed\}$. The operation we made in this small example is defining the *neighborhood* of an element, and more generally, of a set of elements. Thus, we need a theory able to precisely model the *neighborhood* of sets: here we thought of *pretopology* formalism.
Pretopology is a mathematical modeling tool for the concept of *proximity* in mathematical spaces (Belmandt, 1993). It furthermore establishes powerful tools for structure analysis and automatic classification (Le et al., 2008). Pretopology theory generalizes topology theory

and has weaker axioms than topology, allowing us to analyze transformation procedures *step-by-step*, and has interesting properties in discrete case. As pretopology formalism is based on *set theory*, a group of individuals is now considered as a set, allowing us to consider a group as a whole independent entity, which is the main problematic we want to answer.

Pretopological notions

The pretopology formalism comes from classical topology but has weaker axioms. The "closure" process is not necessary idempotent and only the extensivity axiom is required. (Brissaud, 1975) gave the first definition of a pretopological space. (Fréchet, 1928) spaces, (Kuratowski, 1952) closure axioms and (Cech, 1966) closure operator have been preponderant in definition of a pretopological space and of the *pseudo-closure* map (Def. 1 and Fig. 2). Some definitions of the formalism used in pretopology can be found in (Belmandt, 2011).

Definition 1 (Pseudo-closure) *Let consider a non-empty set E, and $\mathcal{P}(E)$ which designates all the subsets of E. A* pseudo-closure *map, denoted by $\boldsymbol{a}(.) : \mathcal{P}(E) \longrightarrow \mathcal{P}(E)$, is a map meeting the following conditions:*

1. $a(\emptyset) = \emptyset$
2. $\forall A \in \mathcal{P}(E), A \subseteq a(A)$

A *pretopological space* is a set E endowed with a pseudo-closure map.

Definition 2 (Pretopological space) *A pretopological space is a couple (E, a) where E is a non-empty set and* a *is a pseudo-closure map $a(.) : \mathcal{P}(E) \rightarrow \mathcal{P}(E)$.*

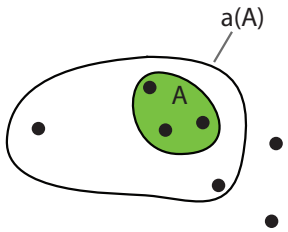


Figure 2: Pseudo-closure of A.

Pseudo-closure operator is useful to model dynamic processes in the considered space (E, a). In social networks, we can say that elements belonging to $a(A)$ are close to A, allowing us to determine direct neighbors. For each pseudo-closure, we absorb new elements ($A \subseteq a^2(A) \subseteq a^3(A) \subseteq ... \subseteq a^k(A)$) which are more and more "distant". Hence, we are able to model complex dynamics like, for instance, information diffusion process. Next, we will briefly give a characterization summarizing the different pretopological spaces that can be encountered. Each type has interesting properties that brings its own corpus of theoretical results and can be used to model particular problems.

Definition 3 (Pretopological space types) *A pretopological space (E, a) can be of various types (considering, when needed, two subsets A and B of $\mathcal{P}(E)$):*

- *a general pretopological space*
- *a $\mathcal{V}$-type pretopological space if:* $A \subseteq B \Rightarrow a(A) \subseteq a(B)$
- *a $\mathcal{V}_D$-type pretopological space if it is a $\mathcal{V}$ type and* $a(A \cup B) = a(A) \cup a(B)$
- *a $\mathcal{V}_S$-type pretopological space if it is a $\mathcal{V}_D$ type and* $a(A) = \bigcup_{x \in A} a(\{x\})$
- *a topological space if it is a $\mathcal{V}_D$-type and it fulfills the following property :* $a(A) = a(a(A))$ *(idempotence)*

It can easily be proven that a pretopological $\mathcal{V}_D$-type structure is a $\mathcal{V}_S$ one in the countable case. Habitually, the most interesting space types are the $\mathcal{V}$-type and the general type, because of their weaker axioms. With these definitions, we are able to model emerging phenomena. We can model more general spaces, with more complex behaviors and more complex *neighborhoods*. To understand this approach, let us return to our network example with the four friends and define the problem by applying pretopology formalism. We build a pretopological space (E, a) with four elements in E and we build $a(A)$ with $A \in \mathcal{P}(E)$ such as:

$a(A) = A \cup$ people who are friends with each person of A.

With the same assumptions, we make the same operations as in the previous example:

- $a(\{John\}) = \{John, Tim, Ben, Ed\}$
- $a(\{Ed\}) = \{Ed, John, Ben\}$
- $a(\{John, Ed\}) = \{John, Ed, Ben\}$ (Fig. 3)

As we notice, $a(\{John, Ed\}) \neq a(\{John\}) \cup a(\{Ed\})$. The space is not a $\mathcal{V}_D$-type. We have also $\{John\} \subseteq \{John, Ed\}$ but not $a(\{John\}) \subseteq a(\{John, Ed\})$. The space is not a $\mathcal{V}$-type either. We observe a general pretopological space. Pretopology helps us to define neighborhoods which are not convenient to describe with other formalisms.

In a network, diffusion process modeled by the pseudo-closure can stop its progression. Such a configuration is called a *closure* (Dalud-Vincent et al., 2009) (Fig. 4).

Definition 4 (Closed subset and Closure) *Let (E, a) a pretopological space:*

- *a subset F of E such as $a(F) = F$ is called a* closed *subset of E for $a(.)$.*

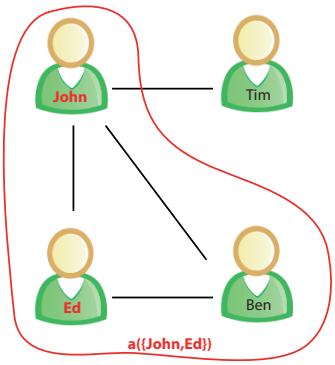


Figure 3: Pseudo-closure of {John,Ed}.

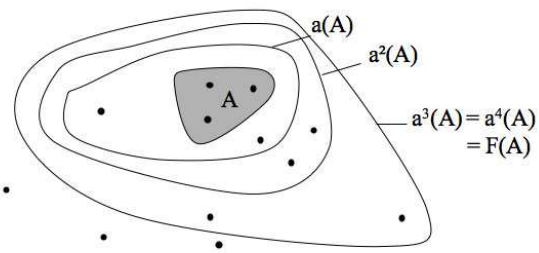


Figure 4: Iterated application of the pseudo-closure map leading to the closure.

- *an* elementary closed subset*, denoted as F_x, is the closure of a one element set $\{x\}$ of E.*

Closure is very important because of the information it gives about the "influence" or "reachability" of a set, meaning, for example, that a set A can influence or reach elements into $F(A)$, but not further.

Modeling group interactions
First, we have to define a social network with pretopology concepts. A social network is a social structure made of nodes (which are generally individuals or organizations) that are tied by one or more specific types of binary or valued relations (Degenne, 2004).
In pretopology, we can generalize this definition by saying that a (social) network is *a family of pretopologies on a given set E* (Fig. 5).

Definition 5 (Network) *Let E be a non-empty set.*
Let I a countable family of indexes.
Let $\{a_i, i \in I\}$ a family of pretopologies on E.
The family of pretopological spaces $\{(E, a_i), i \in I\}$ is a (social) network on E.

Different kinds of relations with different natures can be represented: for instance, we can model a social network containing work colleagues and geographic relations. We can build a pseudo-closure answering to a modeling problem, having for example a person close

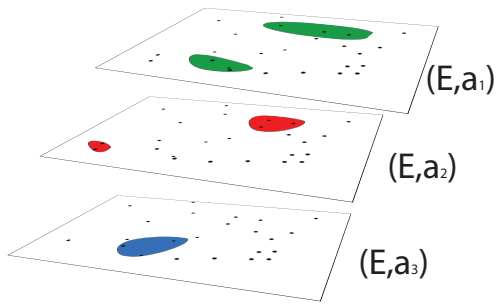


Figure 5: Social network defined with pretopology and composed of three different pseudo-closures.

to another if they are friends, working together, and living close to each other (using a given distance threshold). The following paragraphs give to the reader examples of how we can build pseudo-closures depending on the data or information we want to use for our modeling.

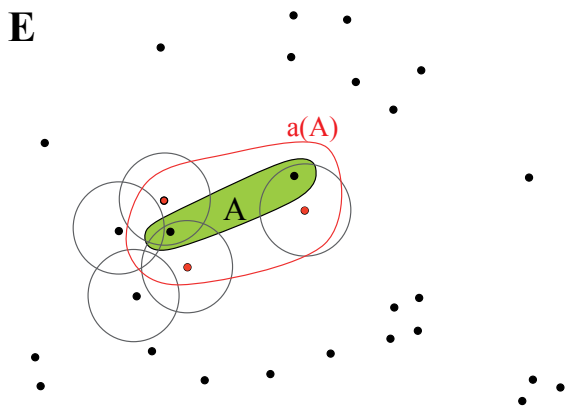
Example 1: Metric Space We represent the relations between people with an euclidean distance (the people are represented as nodes), considering for instance the geographical location of each person.
Let E be endowed with a metric defined by a distance d. Let r be a positive real. For each element x of E, $B(x, r)$ is a ball with the center x and a radius r defined by :

$$B(x, r) = \{y \in E / d(x, y) \leq r\}$$

A pseudo-closure $a(.)$ can be builded with $B(x, r)$:

$$\forall A \in \mathcal{P}(E), a(A) = \{x \in E / B(x, r) \cap A \neq \emptyset\}$$

The pseudo-closure $a(A)$ is composed of all elements of A and all elements $y \notin A$ such that y is within a distance of at most r from at least one element of A (Fig. 6).



$$a(A) = \{ x \in E / B(x,r) \cap A \neq \emptyset \}$$

Figure 6: Pseudo-closure of A in a metric pretopological space.

Example 2: Binary Space In social networks, we are confronted with non-metric relations, representing qualitative information (friendship relation, etc.). In such a space, the elements of E are bound by n reflexive binary relations R_i with $i \in \mathbb{N}^*$. R is not necessarily symmetric. We define $\forall\, i \in \{1, ..., n\}$:

$$R_i(x) = \{y \in E / x R_i y\}$$

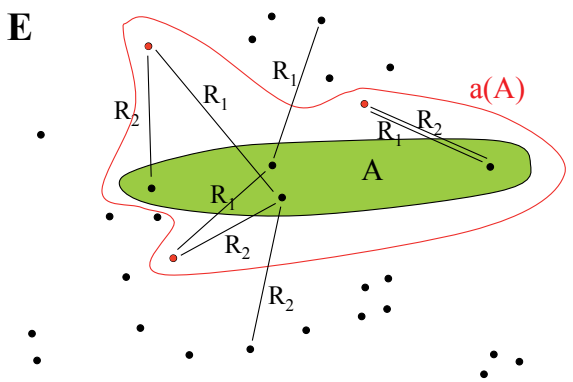
$$R_i^{-1}(x) = \{y \in E / y R_i x\}$$

As the relation R is reflexive, x belongs to $R(x)$ and to $R^{-1}(x)$. We can construct two pseudo-closures:

$$\forall A \in \mathcal{P}(E), a(A) = \{x \in E / \forall i \in \{1, ..., n\}, R_i(x) \cap A \neq \emptyset\}$$

$$\forall A \in \mathcal{P}(E), a(A) = \{x \in E / \forall i \in \{1, ..., n\}, R_i^{-1}(x) \cap A \neq \emptyset\}$$

These pseudo-closures are not equivalent when R is not symmetric. The pseudo-closure $a(A)$ is composed of A (R is reflexive) and of all elements y which have R_n (or R_n^{-1}) relations with, at least, one element of A. An illustration of this kind of space with two binary relations is illustrated in Fig. 7. In this example, an element x of E belongs to pseudo-closure of A if x is in relation with an element of A according to relation R_1 and R_2.



a(A) = { x∈E / ∀i∈{1,2} R_i(x)∩A≠∅ }

Figure 7: Pseudo-closure of A in a binary pretopological space.

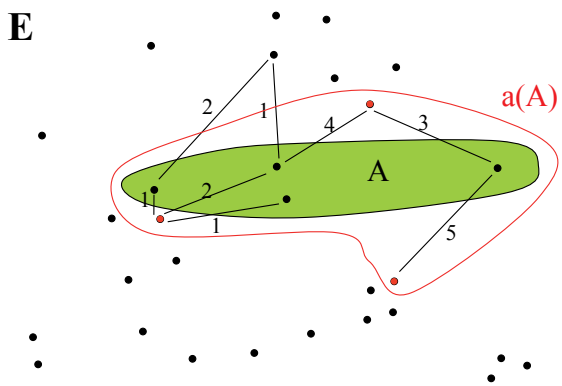
Example 3: Valued space In order to model certain problems, binary relations are not sufficient. We therefore need to have a value (integer, real, function, ...) on the links. In this kind of space, the elements of E are bound by a valued relation. For instance, we can define an integer value v on relations as :

- $E \times E \rightarrow \mathbb{N}$
- $(x, y) \rightarrow v(x, y)$

We can build a pseudo-closure giving different results between singletons and sets ($E - A$ and A^c are equivalent notations corresponding to the complementary of A and s is an integer):

$$\forall A \in \mathcal{P}(E), a(A) = \{y \in E - A \,/\, \sum_{x \in A} v(x, y) \geq s\} \cup A$$

The pseudo-closure $a(A)$ is composed of A and of all elements y where the sum of valued edges between some elements of A and y is greater than the threshold s. Figure 8 gives an illustration of this space with $s = 4$. This kind of modeling can be used in social networks where weighted relations are necessary, and illustrates the interest of the pretopology modeling. Indeed, this example shows that group behavior is different than the "sum" of individuals composing it ($\mathcal{V}$-type). In Fig. 8, the bottom left person (red) is absorbed because he knows 2 persons (value = 1) a little, and better another one (value = 2), so he can be considered as a friend of the group A. If we take each individual of A saying this external individual, the bottom left person, is friend of A if the value of a link is superior to $s = 4$, he will not be taken in account.



a(A) = { y∈E-A / $\sum_{x \in A}$ v(x,y) ≥ 4 }∪A

Figure 8: Pseudo-closure of A in a valued pretopological space.

A Complex Group Interaction Model

We show how we can model social networks with complex interactions. The following example is taken from (Sampson, 1968) study, a "toy" dataset which represents several relations between monks in a cloister, and has been often used in sociology studies. It concerns the social interactions which have been collected as numerous sociometric rankings. A lot of relations are coded, but we focus on two of them: *esteem* and *disesteem*. Relations are non-symmetric and weighted with three integer values, from 1 to 3 (3 indicates the highest or first choice and 1 the last choice). We builded our model on the following assumptions, that a person y is close to another person x if:

- there is an esteem relation from y to x according to a chosen threshold.

- there is no disesteem relation from x to y according to a chosen threshold.

This proximity definition makes sense if we consider group coalitions. We base our model on the following hypothesis: if a person asks other people "who wants to join my group ?", people who have the greatest esteem for him should join his group more than the others. But this person won't accept in his group people he doesn't esteem, even if these people have esteem for him. We set two valued non symmetric relations with E the set containing individuals :

- $E \times E \rightarrow \mathbb{N}$
- $(x, y) \rightarrow est(x, y)$
- $(x, y) \rightarrow disest(x, y)$

with est the esteem relation and $disest$ the disesteem relation.

We build our pseudo-closure as:

$$\forall A \in \mathcal{P}(E), a(A) = \{y \in E - A \,/\, \sum_{x \in A} est(y, x) \geq \alpha \wedge \sum_{x \in A} disest(x, y) < \beta\} \cup A \text{ with } \alpha, \beta \in \mathbb{N}.$$

With $\alpha = 3$ and $\beta = 1$, we put a strong restriction on neighborhood: if we apply pseudo-closure on a set A, people would be in $a(A)$ if they have esteem with one or several persons of A (according to α), and if no people of A have disesteem for them (see Fig. 9.).

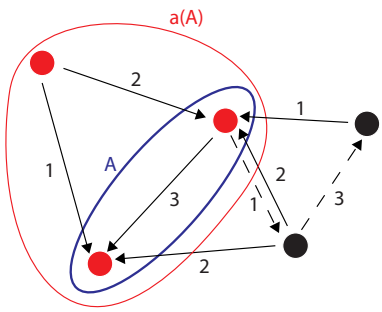


Figure 9: Pseudo-closure of A with $\alpha = 3$, $\beta = 1$. Esteem relation is plain and disesteem relation is dashed.

A typical question (Everett and Borgatti, 1999) in social networks is: who is the person who has the best ability to gather a group around him ? Who is the most influent person ? In this model, finding the largest group following the rules of our neighborhood can be treated by building $\mathcal{F}_e$, the family of all elementary closed subsets of E (see Def. 4.). When closure is applied to each singleton of E, it reveals that `VICTOR_8` is the person who can rally the biggest amount of people in the network (Fig. 10.). Results of our mathematic model has been simulated in JAVA using *PretopoLib* library which implements all pretopological concepts (Levorato and Bui, 2008).

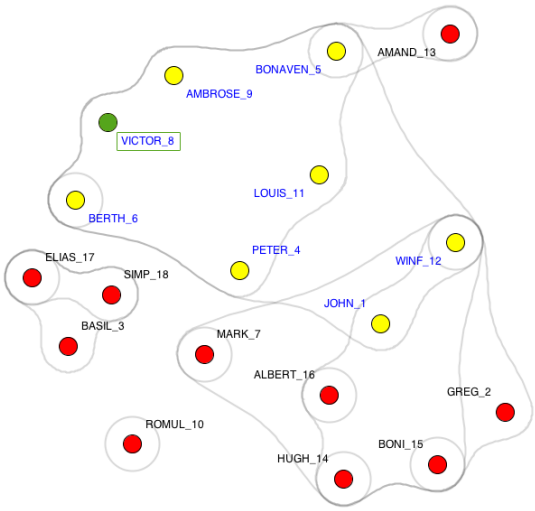


Figure 10: Simulation of our model in order to find most influent person according to our neighborhood definition.

Conclusion

In this work, examples of social networks models based on pretopology formalism have been presented. Building a pseudo-closure map depends on the network nature and problem representation, and can need several pseudo-closures to be combined to obtain specific results. Several previous works have investigated in modeling complex systems with pretopology and have shown some interesting results in: modeling the impact of geographic proximity on scientific collaborations (Largeron and Bonnevay, 2002), structuring and clustering data (Le et al., 2008), modeling pollution phenomena (BenAmor et al., 2008), or analyzing communities of the Web (Levorato and Bui, 2007). We presented ideas and concepts which lead to the following contribution: we proposed a general mathematical framework in order to model groups in social networks. With definitions and examples, we have highlighted the interest to model complex interactions using the pretopology formalism, generalizing graph theory. Furthermore, the introduction of various pseudo-closures functions through examples should be quite helpful in understanding our approach, and we showed an example of how complex group interactions can be modeled.

Future work can be axed on some interesting new methods using generalized measures to find groups which maximize a given centrality (degree, betweeness, closeness), to optimize the "efficiency" of a group, or to identify emergent groups in a network. Moreover, centralities measures could be generalized and based on the same mathematical framework.

REFERENCES

Belmandt, Z. (1993). *Manuel de prétopologie et ses applications*. Hermes Sciences Publications.

Belmandt, Z. T. (2011). *Basics of pretopology*. Hermann.

BenAmor, S., Bonnevay, S., Bui, M., and Lamure, M. (2008). A stochastic and pretopological modeling aerial pollution of an urban area. *Revue Santé et Systémique*, 10(3-4):73–85.

Brissaud, M. (1975). Les espaces prétopologiques. *Compte-rendu de l'Académie des Sciences*, 280(A):705–708.

Cech, E. (1966). *Topological Spaces*. John Wiley and Sons.

Dalud-Vincent, M., Brissaud, M., and Lamure, M. (2009). Closed sets and closures in pretopology. *International Journal of Pure and Applied Mathematics*, 50(3):391–402.

Degenne, A. (2004). Between tool and theory, social networks. *Social Networks of the Internet*.

Everett, M. and Borgatti, S. P. (1999). The centrality of groups and classes. *Journal of Mathematical Sociology*, 23(3):181–201.

Everett, M. and Borgatti, S. P. (2005). Extending centrality. In *Models and methods in social network analysis*, pages 57–76. Cambridge University Press.

Fréchet, M. (1928). *Espaces abstraits*. Hermann.

Kuratowski, C. (1952). *Topologie*. Warszawa.

Largeron, C. and Bonnevay, S. (2002). A pretopological approach for structural analysis. *Information Sciences*, 144:169 – 185.

Le, T. V., Kabachi, N., and Lamure, M. (2008). Pretopology and a homogeneous method for data clustering. In *RIVF'08 conference*. IEEE Computer Press.

Levorato, V. and Bui, M. (2007). Modeling the complex dynamics of distributed communities of the web with pretopology. In *I2CS'07 conference*.

Levorato, V. and Bui, M. (2008). Data structures and algorithms for pretopology: the java based software library pretopolib. In *I2CS*, pages 122–134, Schoelcher, Martinique. IEEE.

Newman, M., Barabási, A.-L., and Watts, D. J. (2006). *The Structure and Dynamics of Networks*. Princeton University Press.

Saha, B. and Getoor, L. (2008). Group proximity measure for recommending groups in online social networks. In *The second SNA-KDD Workshop in ACM SIGKDD*. ACM Press.

Sampson, S. (1968). *A Novitiate in a Period of Change. An Experimental and Case Study of Social Relationships.* PhD thesis, Cornell University.