Homophily and Structure in Multiplex Networks

Basit Chaudhry	Chris Marton	Hana Shepherd
RAND and UCLA	M.I.T.	Princeton University
Los Angeles, CA, USA	Cambridge, MA, USA	Princeton, NJ, USA

Abstract

A large body of work attests to the prevalence of homophily—similarity between individuals—in social networks, though which mechanisms are responsible for the prevalence of homophily remains subject to debate. Additionally, there is a paucity of work examining the effect of multiple types of ties on social influence processes and network homophily generally. This project explored the dynamics of homophily within a multiplex network and the effect of social influence on network structure itself.

1 Introduction

1.1 Network Homophily and Generating Mechanisms

Empirical work on the prevalence of "homophily", or similarity between individuals, within social networks has demonstrated that it is one of the most robust and pronounced characteristics of social networks (McPherson, Smith-Lovin, & Cook 2001). Of course, there are many possible criteria on which individuals could be determined to be "similar" to one another. In their seminal work on homophily in networks, Lazarsfeld & Merton (1954) distinguished between *status homophily*, whereby individuals are considered similar to one another on the basis of informal, formal or ascribed status under which demographic variables such as age, gender, and race fall, and *value homophily*, whereby individuals are considered similar to one another on the basis of shared values, attitudes, and beliefs. There are three possible means by which networks could become homophilous: 1. Individuals who are similar to one another are more likely to form ties such that there is strong selection into networks where individuals have the same characteristics, 2. individuals who are tied to one another are more likely to influence each other; this is especially true of value homophily for the obvious reasons that most forms of status are subject to social processes at levels beyond those involved in direct social interactions between people or small groups of people, or 3. some combination of each of these occurs where there is a feedback processes between the characteristics of individuals within a network and the basic structure of that network. The evidence for which of these processes is the most appropriate for describing the creation of homophilous networks is not conclusive as some bodies of literature suggest selection processes are most responsible for homophilous networks while others suggest that social influence plays a substantial role.

From empirical social network findings such as those by Kandel (1978) and Cohen (1977) who studied adolescent friendship networks, McPherson, Smith-Lovin, & Cook (2001) conclude that "the selection into relationships with similar others appears to be a much more powerful force than interpersonal influence within the friendship network" and that more generally "selection almost certainly trumps influence or attrition" (429).

This, however, seems to run counter to a large literature, particularly within social psychology, regarding how individuals' attitudes, beliefs and behaviors are profoundly influenced by those around them (e.g. Asch 1955, Festinger 1954, Sherif & Hovland 1961). These classical findings in emphasize the susceptibility of beliefs and attitudes to the influence of others, suggesting that the emergence of value homophily through social influence processes is quite feasible. In addition, there is evidence that those in networks where their own attitudes match those of other network members are more likely to maintain those attitudes. Visser & Mirabile (2004) experimentally manipulated the similarity of attitudes within a network and found that individuals in networks where others held attitudes similar to their own were more likely to be resistant to changing their attitude than were individuals in networks with heterogenous attitudes among members. Other work from sociology on the diffusion of ideas and learning suggests that interactions between individuals changes the "underlying sociocultural environment" where interactions lead to the formation of subgroups, subcultures, and the dominance of particular beliefs (see Carley 2001). These findings suggest that "group members become more alike and their attitudes and beliefs become correlated" (Carley 2001). Carley (2001) additionally reviews a number of studies which assume that interactions between individuals and the exchange of information, ideas, attitudes and beliefs within those interactions changes the structure of relations and thus the network structure itself. An article by Marsden & Friedkin (1993) on social influence in networks also assumes the influence of social relations within a network on attitudes and behaviors of individuals. Axelrod (1997) takes as a given that "similarity leads to interaction, and interaction leads to still more similarity". This evidence seems to point towards a seminal role of social influence and subsequent network change in shaping the composition of networks.

1.2 Multiplex Networks

Networks containing multiple types of edges, multiple types of relations between actors, have long been important to social scientists given that empirical networks are obviously composed of many different types of ties (for example, see Fischer 1982). Much of the social scientific work on networks with multiple types of ties (also referred to as multiplexity) has focused on developing metrics and methods of analyzing such data (e.g. Minor 1983, Fienberg, Meyer, & Wasserman 1985, Krohn, Massey, & Zielinski 1988). The interest in incorporating multiplexity into accounts of homophily in networks stems from the basic observation that the flow of influence across networks might depend on the nature of particular *types* of ties between actors[1]. For example, social influence with regards to health-related behaviors might spread based on familial or friendship ties more than it might spread through work ties. Much work on the diffusion and influence across networks does not take into account how the *types* of ties and importantly how these ties

[1] In addition to this observation, McPherson, Smith-Lovin, & Cook (2001), in their call for research into the effects of multiplexity on homophily, note that "if different types of relations are structured by different levels of homophily on different dimensions, then multiplex relations among individuals may create systematic, important patterns of cross-cutting social circles. Attention to this complexity may produce findings as important for the larger issue of the integration of society as did Peter Blau's (1977) groundbreaking insights about that impact of consolidated (correlated) dimensions."

are related structurally within a network with *multiple* types of ties, might affect social influence processes and the degree of homophily within a network.

Empirically, homophily in networks has been demonstrated across many types of relations between individuals including friendship, advice, mere contact and "knowing", marriage and more. Networks are commonly homogenous with regard to demographic characteristics of gender, age, race, religion, social class and with regard to characteristics such as network position, behavior and attitudes (McPherson, Smith-Lovin, & Cook 2001). Some evidence suggests that as the number of types of relationships between individuals increases, homophily increases "indicating that homophily on each type of relation cumulates to generate greater homophily for multiplex than simplex ties" (McPherson, Smith-Lovin, & Cook 2001).

1.3 Model Purpose

This project sought to examine the dynamics of homophily in social networks using agent-based modeling techniques. The main features of interest from the social networks literature were homophily between individuals (this project is primarily concerned with similarity between individuals based on value homophily) and how homophily leads to the creation of social ties, social influence processes between interacting individuals, dissolution of ties between dissimilar individuals, and the presence of multiple types of ties between individuals. There were two substantive questions the model was designed to examine: 1. How does homophily impact network structure and how is it impacted by network structure over time? Can relatively stable levels of network homophily be reached and under what conditions does this occur? How does the inclusion of agents who are not homophilous and not incorporated into the already-existing network structure influence network structure and homophily? 2. What effect does having multiple types of ties have on both the social influence processes that are central to the creation and maintenance of homophilous social networks and on the network structure itself? Is the effect of multiple types of ties on network structure always mediated through their effect on social influence processes? Agent-based modeling techniques are well-suited for these questions, given that they provide a method for systematically evaluating the impact of various factors on network homophily and structure. Our model is also able to compare the relative effects of selection into homophilous networks to social influence within networks creating greater degrees of homophily, on overall network homophily and structure, though the findings are not reported in this paper.

As outlined in greater detail below, the model contains two types of agents, one where relations of two different types already exist between agents, and another where there are no connections between agents of that type or agents who are already in a network. They new agents are intended to disrupt the homophily of the existing network. Each agent has a bit string, which represents the "values" or beliefs or other characteristics by which agents can be similar to one another. Part of the bit string corresponds to similarity related to one type of network tie and the other part corresponds to similarity related to the other type of network tie[2]. Agents interact with others and are influenced by the

[2] While we are interested in the effect of multiple types of *ties* on social influence and structure, other modeling work exploring the effect of multiple *features* or characteristics of agents on social influence is

characteristics of each other represented by the bit strings and form or dissolve ties on the basis of similarity or difference in the parts of the bit strings corresponding to the two types of ties[3]. The parameters of particular interest for the questions posed are the social influence thresholds, or how much similarity there needs to be between two interacting agents for them to influence one another and for one of the agents to change, and tie making and breaking thresholds, which determines how much similarity there needs to be between two agents to either make a tie or to break an existing tie.

The interest in the effect of multiple types of ties on network homophily stems from the assumption that multiple types of ties change both aspects of the process by which networks might become homophilous; they change both how similarity between individuals shapes network structure and how network structure shapes similarity between individuals.

Section 2 of this paper reviews the specifics of the implemented agent-based model. Section 3 reviews some of the results of a preliminary analysis of experiments with the model. Future work will focus on developing experiments with the model to examine the effect of the selection vs. social influence mechanisms proposed in the literature on homophily in the network and network structure.

2 Model

2.1 Algorithm Outline

The impact of multiple types of network ties was represented by a population of N agents capable of forming two types of ties, termed types 1 and 2. Each agent was characterized by a bit string of length L, where L is an even integer, which is referred to as a logic string. Each element of the logic string was allowed to take a value of either 0 or 1.

At each time step, the agents are able to make ties of each type with "neighboring" agents, depending on the similarity of their respective logic strings. The agents are also able to break ties with agents to which they had a previous link depending on the similarity of their respective logic strings. Agents are also able to compare and synchronize their logic strings with agents to which they are currently connected in the network.

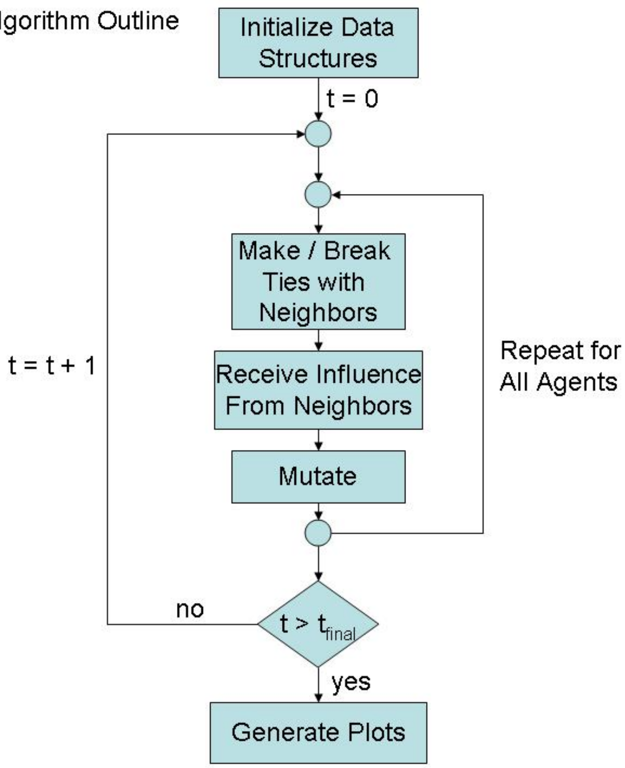
---

relevant. For example, work by Axelrod (1997) has explored the effect of having multiple features of agents, where agents influence one another based on similarity of features, on global group properties. In his model, based on the rule that the more similar an agents' neighbor, the more likely that agent is to adopt one of their traits, he finds that polarization of the group as a whole emerges and that the number of homogenous regions decreases with the number of features of agents that are included in the model.

[3] This feature of the model is empirically supported by a number of studies which report that individuals who are dissimilar are more likely to dissolve their ties to one another (Hallinan & Williams 1989, Tuma & Hallinan 1979, Popielarz & McPherson 1995, Burt 2000).

The agents are initialized with a random network structure and random logic strings. The agents are then allowed to develop the networks and alter their logic strings according to the system rules. When the system has run for the specified length of time, several metrics can be used to determine the structure of the two networks, and the characteristics of the logic strings of the agents. The algorithm is also designed to allow for an exogenous shock of new agents to be introduced after some length of time, allowing for a more dynamic analysis of network structure and the change in logic strings. This outline is illustrated in figure 1.

Figure 1: Algorithm Outline



The algorithm was implemented using Matlab 7.0.

2.2 Agent and Population Properties

The network of N agents is described by two adjacency matrices of N by N dimension, **A** and **B**, which identify the ties of each type between agents. Elements of matrix **A** take binary values and indicate ties of type 1 between two agents. For example, if the element in the i^{th} row and j^{th} column of matrix **A** takes a value of 1 it indicates that agents i and j share a tie of type 1. If the same element of **A** assumed a value of 0, then it would indicate that agents i and j did not share a tie of type 1. The matrix **B** uses the same convention to describe ties of type 2. The **A** and **B** matrices are each symmetric, and the diagonal elements are required to take values of 0 as self-ties are not allowed.

As mentioned above, each agent is characterized be a logic string of length L. The first half of the logic string is related to network ties of type 1, and the second half is related to type 2. For example, when agents look to make a tie of type 1 with another agent, the similarity measure used to determine whether or not the tie can be formed is based on the number of logic bits in the first half of the two logic strings that contain the same value in the same position. Also, two agents that share a tie of type 2 only are more likely to homogenize their logic bits on the second half of their logic strings.

2.3 Population Interaction

At each time step, each agent performs the same three tasks: make and break network ties, receive influence from tied agents, and possibly mutate the logic string.

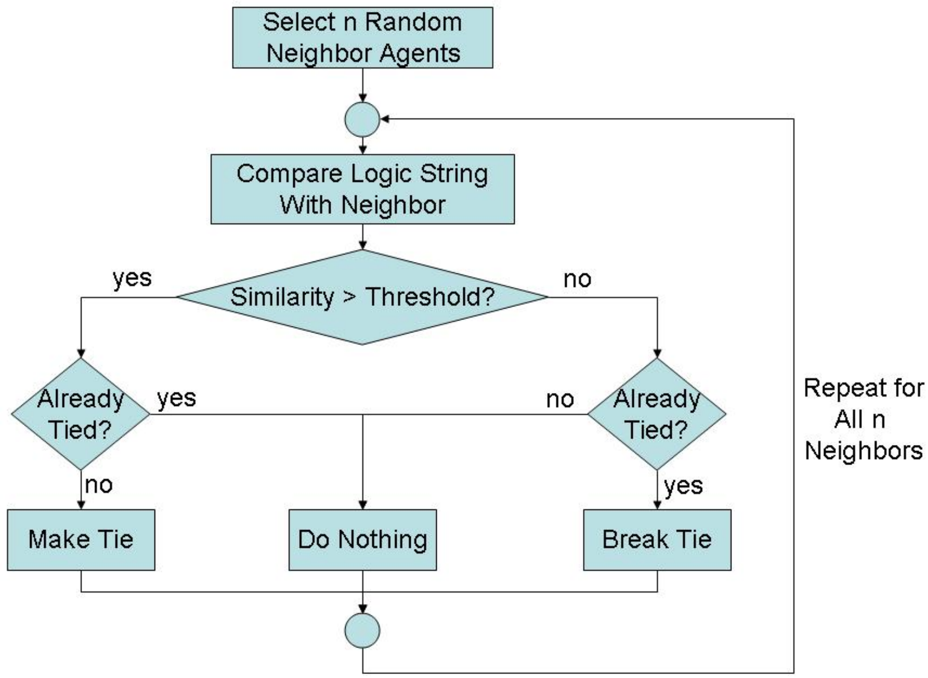
1) Make and Break Network Ties

The agent randomly selects *n* other agents randomly from the population (with uniform probability) and classifies these agents as in its "neighborhood". The procedure for type 1 network ties is as follows.

For each agent in its neighborhood, the current agent calculates the number of shared logic bits between itself and the neighboring agent, over the first half of the logic string. If the agent is already type-1 tied to the neighboring agent in question, the similarity is compared to a break threshold. If the similarity is below this break threshold then the tie is broken; otherwise, there is no change to the adjacency matrix **A**. If the agent is not type-1 tied to the neighboring agent, then the similarity is compared to the make threshold. If the similarity is above this make threshold then the two agents make a type 1 tie; otherwise, no change is made to the adjacency matrix **A**.

The same procedure is followed for network ties of type 2. The procedure is illustrated in figure 2.

Figure 2: Make and Break Ties

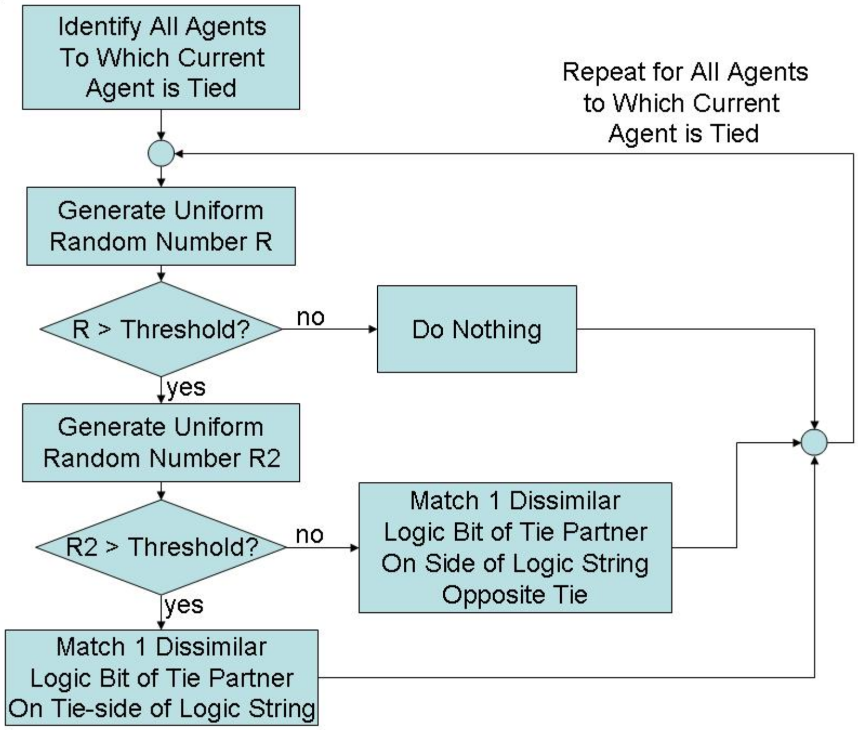


2) Receive Influence from Tied Agents

The current agent first identifies all the agents with which it shares type 1 network ties. The agent then generates a uniform random number between 0 and 1 with an influence threshold. If the random number is greater than the influence threshold, then the agent moves on to the next agent with which it shares a type 1 network tie. However, if the random number is less than the influence threshold, then another uniform random number between 0 and 1 is generated, and compared to a same side threshold. If the random number is less than the same side threshold, then a logic bit on the type 1 network side of the logic string that is dissimilar between the two agents is selected, and the logic bit of the current agent is flipped to match that of the agent with which it is tied. Otherwise the logic bit to be flipped is selected from the half of the logic string not associated with type 1 ties. The current agent then moves on to the next agent with which it shares a type 1 network tie.

The same procedure is repeated for the type 2 network. The procedure is illustrated in figure 3.

Figure 3: Receive Influence



3) Mutation

For each agent, a uniformly generated random number between 0 and 1 is generated and compared to a mutation threshold. If the random number is less than the mutation threshold then a randomly selected logic bit is flipped (i.e. 0 is replaced with 1 or vice versa). If the random number is greater than the mutation threshold then no change is made.

3 Preliminary Results

For this project, we ran a series of experiments varying values of three parameters: the probability of social influence (how likely the bit string of an agent is to be influenced by the agent with which it is interacting), the threshold for similarity of bit strings required to make a tie with a neighboring agent not already tied and the threshold for difference of bit strings required to break a tie with a neighboring, already tied agent. Again, half of a bit string of an agent corresponded to one type of tie and the other half of the bit string corresponded to the other type of tie. The results presented here examine the distribution of the Hamming distances between any agent and all other agents to which they are tied, where Hamming distance in this case is the number of bits that would need to be switched for the bit strings of two agents to be identical. Thus, a large number of low Hamming distances represents a high degree of similarity or homophily among agents who are tied to one another. The preliminary results reported here on the resulting distributions of Hamming distance in the network refer only to distance calculated on the tie-relevant side (i.e. Hamming distance is measured only with respect to the side of the bit string that corresponds to the same tie). Future work will examine the effect of different parameters on the Hamming distances between bit sequences on *opposite* sides of the strings (i.e. the distance between a sequence that refers to tie 1 for one agent and to tie 2 for the other agent) to explore the patterns of similarity or homophily across different types of ties.

We examine four patterns based on these experiments. In the first two patterns, the threshold for similarity of bit strings between interacting agents that was required to create a tie between those agents was set high, resulting in relatively sparsely connected networks (as evidenced by the differences in the relative frequencies between figures 4 & 5 and figures 6 & 7 below) . For patterns three and four, the threshold for similarity of bit strings between interacting agents that was required to create a tie between those agents was set low, resulting in relatively densely connected networks. The probability of social influence was allowed to vary across a range of parameter values for these scenarios. To model a high threshold for tie formation, we set the similarity between bit strings to nine out of ten. The low threshold for tie formation was set to six out of ten bits. The low threshold for tie breaking was set to three out of ten bits (i.e. if two agents had a distance of three, two or one, the tie between them was broken) and the high threshold for breaking ties was set to five out of ten bits.

To investigate the potential role of social influence in tie formation with respect to value homophily, we varied the probability that an agent is influenced by the bit string of its interacting agent from 0.1 to 0.7 in increments of 0.2. Social influence was found to have little impact on the distribution of Hamming distances within the network when compared with the effects of varying the thresholds for making and breaking network ties. Given these results, we focus the rest of our discussion on the effects of varying tie thresholds.

Scenario 1: High threshold for making ties and low threshold for breaking ties (Figure 4)

In scenario 1, the distribution of Hamming distances between tied agents in the model was skewed rightward. There was a predominance of very low Hamming distances in the model, suggesting that the majority of agents had bit strings that were very similar to those of other tied agents. Homogeneity was demonstrated early on in the time progression of the model. As the simulation proceeded, homogeneity increased.

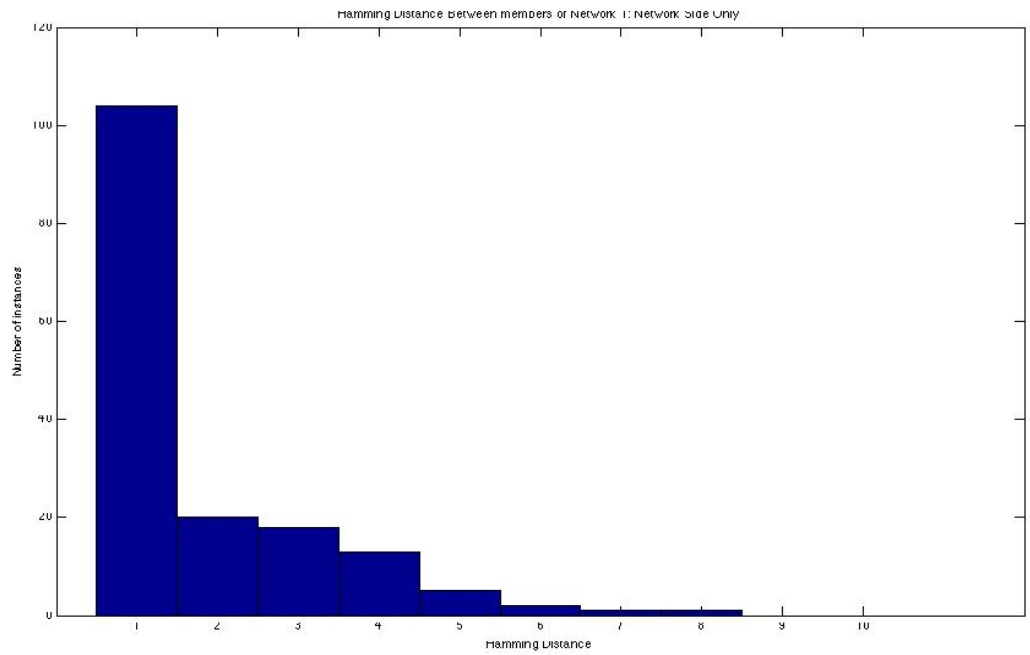


Figure 4: High threshold for making ties, low threshold for breaking ties.

Scenario 2: High threshold for making ties and high threshold for breaking ties (Figure 5)

The results for scenario 2 were similar to those for scenario 1, with significant rightward skewing of the distribution of Hamming distances. Again, there was substantial similarity between the bit strings of most agents in the model. As time proceeded in the model, homogeneity increased slightly faster than in scenario 1.

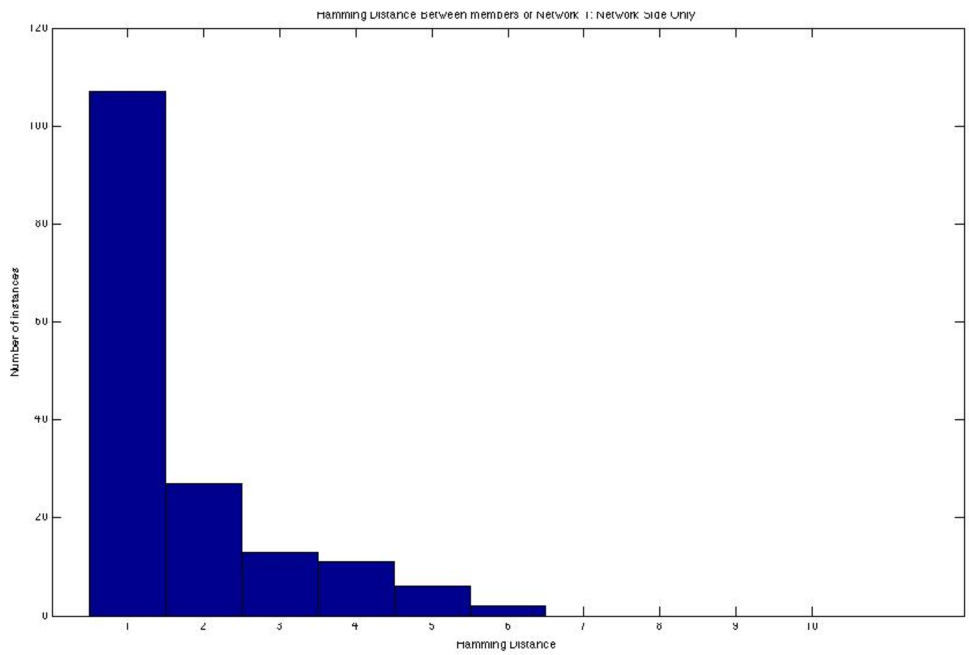


Figure 5: High threshold for making ties, high threshold for breaking ties.

Scenario 3: Low threshold for making ties and low threshold for breaking ties (Figure 6)

With the lower thresholds for making ties resulting in more densely connected networks (see count for number of instances below compared to those in figures 4 & 5), considerably more heterogeneity among the bit strings of agents was observed.

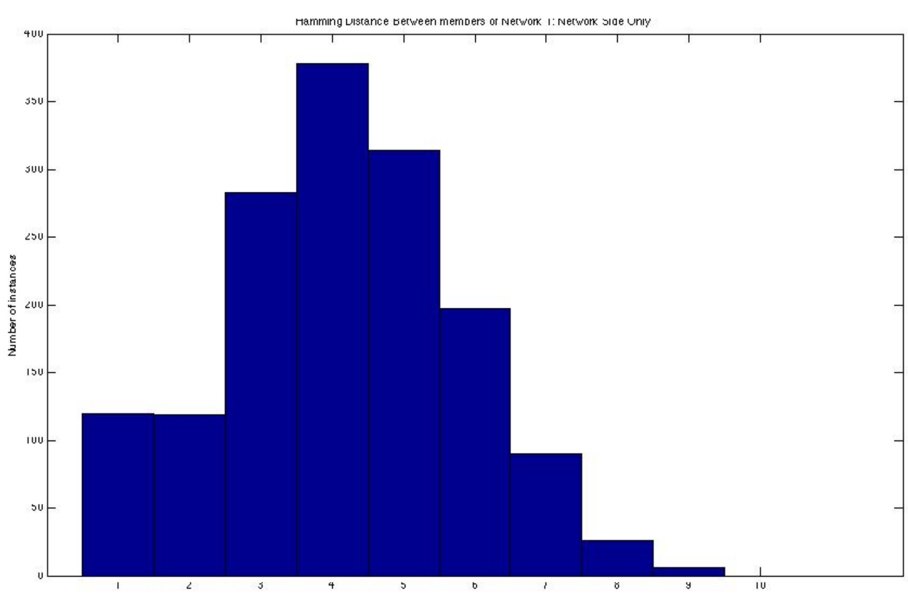


Figure 6: Low threshold for making ties, low threshold for breaking ties.

Scenario 4: Low threshold for making ties and high threshold for breaking ties (Figure 7)

Results for scenario 4 were similar to those for scenario 3, where heterogeneity among agents indicated by the prevalence of higher Hamming distances, was more common than in scenarios 1 or 2.

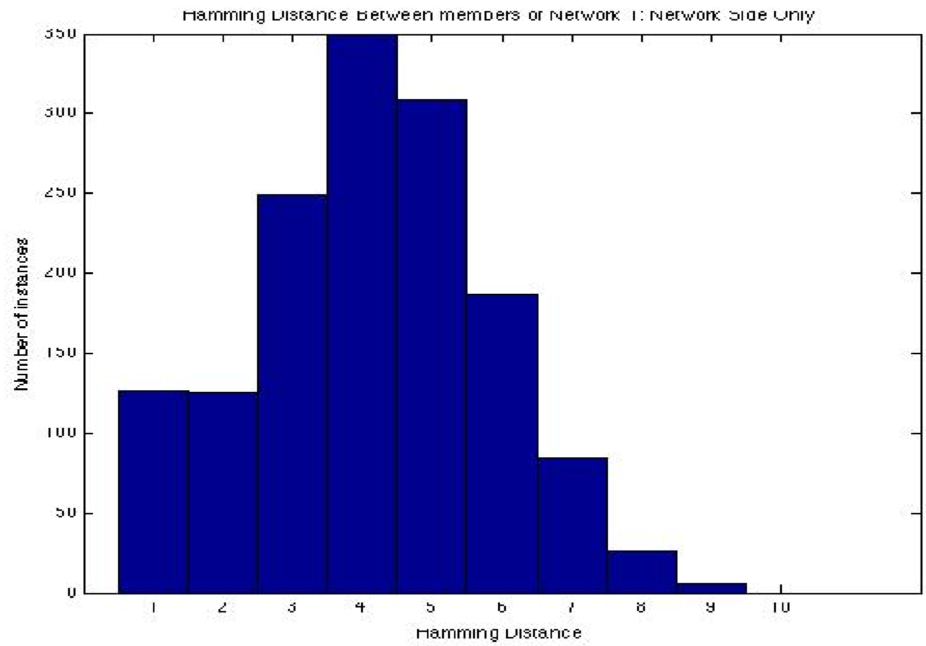


Figure 7: Low threshold for making ties, high threshold for breaking ties.

4 Conclusions

Preliminary results from our model suggest that the similarity among agents or the "value homophily" as represented by Hamming distances in the bit strings is primarily dependent on the threshold levels for tie making and tie breaking, rather than on the probability of being influenced by an interacting agent. Further explorations of parameter space will more rigorously test this initial observation to determine under what conditions the probability of social influence might impact the similarity among agents. We observe similar patterns in the distribution of the Hamming distances when the threshold for making ties is high or low, even if the threshold for breaking ties is varied. High thresholds for making ties, resulting in sparsely connected networks (or networks with a low average degree distribution) seem to produce the greatest similarity among agents. Low thresholds for making ties, resulting in densely connected network structure, produce greater heterogeneity among agents and greater variance across agents in the average Hamming distances.

Our research has several important limitations that we hope to build on in future research. First, due to time constraints, our analysis of the effects of social influence and tie thresholds were limited to the four scenarios described above. There are many other issues to explore regarding this model. Here we limited our outcome metrics to Hamming distances examining the similarity among agents but we would like to analyze similarity among agents based on the structure of the network, i.e. to examine the average Hamming distance distribution in clusters of the network to develop a more nuanced picture of value homophily within the network. Additionally, the results presented here did not examine the effect of the new agents in the model and particularly their equilibrium average degree and structural position within the network. Most importantly for the substantive questions regarding the emergence of network homophily, future work will need to explore the conditions under which social influence creates value homophily as opposed to selective tie-formation to already-similar agents. Second, given the relatively paucity of research on multiplex networks we chose a parsimonious model for our agent based simulation. Our goal was to spur hypothesis generation and theory development rather than making definitive conclusions. We hope to further refine the model with respect to constructs relevant to network multiplexity and value homophily. Finally, our model sets initial network structure and agent logic strings at random. While these assumptions may hold for some types of networks, they may not for many other types of networks. We hope to address this limitation by exploring parameters regarding initial network conditions that will make our model valid for a wider range of conditions.

5 Acknowledgements

We would like to thank the Santa Fe Institute and Dan Rockmore and the other coordinators of the 2006 Complex Systems Summer School.

References

Asch S. 1955. Opinions and social pressure. *Scientific American* 193: 31-35.

Axelrod R. 1997. The Dissemination of culture: A Model with local convergence and global polarization. *The Journal of Conflict Resolution* 41(2): 203-206.

Burt R. 2000. Decay functions. *Social Networks* 22: 1-28.

Carley K. 2001. Learning and using new ideas: A Sociocognitive perspective. In *Diffusion Processes and Fertility Transition,* ed. J. Casterline, pp. 179-207. Washington D.C.: National Academy Press.

Cohen J. 1977. Sources of peer group homogeneity. *Sociology of Education* 50: 227-241.

Fienberg S, Meyer M, Wasserman S. 1985. Statistical analysis of multiple sociometric relations. *Journal of the American Statistical Association* 80: 51-67.

Fischer C. 1982. *To Dwell Among Friends.* Chicago: University of Chicago Press.

Hallinan M, Williams R. 1989. Interracial friendship choice in secondary schools. *American Sociological Review* 54: 67-78.

Kandel D.B. 1978. Homophily, selection and socialization in adolescent friendships. *American Journal of Sociology* 84: 427-436.

Krohn M, Massey J, Zielinski M. 1988. Role overlap, network multiplexity, and adolescent deviant behavior. *Social Psychology Quarterly* 51: 346-356.

Lazarsfeld P, Merton R. 1954. Friendship as a social process: A Substantive and methodological analysis. In *Freedom and Control in Modern Society,* ed. M. Berger, pp. 18-66. New York: Van Nostrand.

Marsden P, Friedkin N. 1993. Network studies of social influence. *Sociological Methods and Research* 22: 127-151.

McPherson M, Smith-Lovin L, Cook J. 2001. Birds of a feather: Homophily in social networks. *Annual Review of Sociology* 27: 415-444.

Minor M. 1983. New directions in multiplexity analysis. In *Applied Network Analysis*, eds. R. Burt, M. Beverly Hills: Sage Publications.

Popielarz P, McPherson M. 1995. On the edge or in between: Niche position, niche overlap, and the duration of voluntary memberships. *American Journal of Sociology* 101: 698-720.

Sherif M, Hovland C. 1961. *Social judgment: Assimilation and contrast effects in communication and attitude change.* New Haven: Yale University Press.

Tuma N, Hallinan M. 1979. The effects of sex, race and achievement on schoolchildren's friendships. *Social Forces* 57: 1265-1285.

Visser PS, Mirabile RR. 2004. Attitudes in the social context: The impact of social network composition on individual-level attitude strength. *Journal of Personality and Social Psychology* 87: 779-795.

Wasserman S. 1987. Conformity of two sociometric relations. *Psychometrika* 52: 3-18.