

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Propagation of Innovations in Networked Groups

Permalink

<https://escholarship.org/uc/item/6rc4q6xq>

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 27(27)

ISSN

1069-7977

Authors

Goldstone, Robert L.

Jones, Andy

Mason, Winter A,

Publication Date

2005

Peer reviewed

Propagation of Innovations in Networked Groups

Winter A. Mason (wimason@indiana.edu)

Department of Psychology, 1101 E. 10th St.
Bloomington, IN 47408 USA

Andy Jones (andjones@cs.indiana.edu)

3406 S. Oaklawn Circle
Bloomington, IN 47401 USA

Robert L. Goldstone (rgoldsto@indiana.edu)

Department of Psychology, 1101 E. 10th St.
Bloomington, IN 47408 USA

Abstract

A novel paradigm was developed to study the behavior of groups of networked humans searching a problem space. We examined how different network structures affect the diffusion of information about good solutions. Participants made numerical guesses and received scores that were also made available to their neighbors in the network. When the problem space was monotonic and had only one optimal solution, groups were fastest at finding the solution when all of the groups' information was presented to them. However, when there were good but suboptimal solutions (i.e., local maxima), the group connected via a small-world network (Watts & Strogatz, 1998) was faster at finding the best solution than all other network structures.

Keywords: Social networks; group behavior; small-worlds; information diffusion; innovation diffusion

Background

Humans are uniquely adept at adopting each other's innovations. Cultural identity is largely due to the dissemination of concepts, beliefs, and artifacts across people. Our capacity for imitation has been termed "no-trial learning" by Bandura (1965), who stressed that people perform behaviors that they would not have otherwise considered by imitating one another. While imitation is commonly thought to be the last resort for dull and dim-witted individuals, cases of true imitation are rare among non-human animals (Blackmore, 1999), requiring complex cognitive processes of perception, analogical reasoning, and action preparation. When combined with variation and adaptation based upon reinforcement, imitation is one of the most powerful methods for quick and effective learning.

One early line of research on when people imitate others focuses on conformity in social groups. To some degree, conformity is found when people desire to obtain social approval from others. For example, when people give their answers privately, they are less likely to conform to the group's opinion than when responding publicly (Deutsch & Gerard, 1955). However, the conformity sometimes runs

deeper than this, and people continue to conform to the group's opinion even privately (Sherif, 1935).

Much of the work on conformity focuses on opinions, information that does not have a verifiable component in the world. In other cases, however, information obtained from others is actually right or wrong. When a person discovers a truly better solution to a problem, this innovation spreads through a population just like any kind of information, including opinions. Innovations are especially like opinions when it is difficult to determine if the innovation is better than current practice or if there is no inherent difference, or when the benefits of adopting an innovation are largely due to others using it (e.g., Macintosh vs. IBM or BetaMax vs. VHS).

The choice between relying on information from others and obtaining information on one's own must be made often in everyday life. Seeking out information on one's own requires time and energy, but is often more trustworthy and individually tailored than information learned by word-of-mouth. On the other hand, choosing to use information provided by others can be cost-effective, especially if past experience suggests that the source is reliable. This attractiveness of exploiting information from others can have an impact on the population, because in cases like these, the "I'll do it if you do it" mentality can lead to "tipping points," (Gladwell, 2000) in which a small number of people initiate a positive feedback cycle, leading to an exponential increase in the number of users for a period.

Banerjee (1992) modeled situations where information is collected and distributed sequentially and showed that when the behavior of other agents was considered equally informative as personally obtained information, rational agents repeated the best solution found by the first few agents regardless of the information they obtained themselves. Bikhchandani, Hirshleifer, and Welch (1992) called this an "information cascade" and suggested this process could be the cause of fads, fashion, and other cultural phenomena.

Valente (1996) looked at individuals with different "adopter thresholds" in the context of their social network. This supported work by Granovetter (1978) who first suggested that people act as though they have a threshold

Table 1: Actual average geodesic path length, clustering coefficient, and closeness centrality for the networks used in the experiment.

	Path Length	Clustering	Closeness
Full	1.00	1.00	0.00
Lattice	3.08	0.35	0.014
Small-World	2.61	0.09	0.026
Random	2.57	0.37	0.050

number of friends (or neighbors) that must adopt a solution before they will also adopt the solution. He found that the people who were early in adopting a solution (those with a low threshold) were most influential in causing bandwagoning in a population. Michael Chwe (1999) extended this threshold model and found that the network position of an individual could be more important than their threshold with respect to causing other people to revolt (or adopt an innovation, depending on your chosen metaphor). This highlights the importance of another factor in the diffusion of innovation, the social network structure.

The properties of network topologies have been studied in many different arenas, including neural networks, actor collaboration networks, power grids (Watts & Strogatz, 1998), citation links (Newman, 2001), metabolic networks (Jeong, Tombor, Albert, Oltvai, & Barabasi, 2000), Web links (Albert, Jeong, & Barabasi, 1999), and many more. A wide range of statistics has been developed to describe the global properties of these networks. These properties are usually defined in terms of the nodes, which are the units or individuals in a network, and edges, the connections between them.

First, the *degree* of a node is the number of edges connecting that node to other nodes. The degree of a network is the average degree of all nodes. Second, the *average geodesic path length* is the smallest number of nodes a message needs to go through to link two nodes, averaged across all pairs of nodes. This property has been popularized as the notion of “six degrees of separation” connecting any two people in the world, and has been experimentally supported (Milgram, 1967). The *clustering coefficient* is the proportion of directly connected neighbors of a node that are themselves directly connected with each other, which can be thought of as the “cliquishness” of a network. The *closeness centrality* measure of a network, developed by L. C. Freeman (1979), amalgamates the closeness of all of the nodes in a network into a single measure. The closeness of a node is the inverse of the sum of the geodesic path lengths to all other nodes, which means that on average, a message will reach a node faster if it has a greater closeness. The actual values of these measures for the networks used in the experiment are listed in Table 1.

Erdős & Rényi (1959) were first to thoroughly describe the properties of random networks, in which edges between nodes are generated such that node i and node j are connected with some probability p . These networks tend to have a small average geodesic path length. More formally,

the average path length connecting two randomly selected nodes in a random network is $\ln(N)/\ln(K)$ where N is the number of nodes and K is the degree of each node. With a large number of nodes random networks tend to have a small clustering coefficient, although with fewer nodes the probability of three nodes forming a triangle is higher, and so the clustering coefficient tends to be higher. However, even with fewer nodes they have a greater closeness to each other on average than other networks.

Another useful network structure is a completely regular network, such as a lattice or ring, in which the structure of edges defines a spatial structure solely made up of local connections. In regular lattices, the average path required to connect two individuals requires going through $N/2K$ other individuals. Thus, the paths connecting people are much longer, on average, for lattice than random networks. Additionally, in these networks, the clustering coefficient tends to be high, since nodes that are spatially close tend to be connected to each other, and the closeness tends to be low, because nodes are “farther” from each other on average.

Watts and Strogatz (1998) demonstrated that by starting with a regular structure such as a lattice and randomly rewiring a small number of connections, the resulting “small-world” network still maintains a sort of regular structure because nodes that are connected to the same node tend to be spatially close themselves (and so have a greater closeness), but also has a small average geodesic path length. From an information processing perspective, then, these are attractive networks because the spatial structure of the networks allows information search to proceed systematically, and the short-cut paths allow the search to proceed quickly (Kleinberg, 2000).

Allen Wilhite (2001) compared market trading over various network structures. In one condition, all agents were allowed to trade with any other agent. In another, agents could only trade locally, in small cliques. In a third condition, most agents could only trade locally, but a few could trade globally (i.e., outside of the local clique). In this small-world network, the market reached Pareto equilibrium (the state where no more trades that mutually benefit both traders can be made) even faster than the condition where everyone could trade with everyone. This is further evidence that small-world networks have unique features in the dissemination of information.

There is excellent work studying the diffusion of innovation in real groups (e.g., Ryan & Gross, 1943; Rogers, 1962, 1995), social psychological research on how individuals use information provided by others (Sherif, 1935; Cialdini & Goldstein, 2004), as well as computational models of information transmission (Nowak, Szamrej, & Latané, 1990; Axelrod, 1997; Kennedy, Eberhart, & Shi, 2001). The study reported in this paper tie together these diverse areas by exploring the diffusion of innovative ideas among a group of networked participants, each of who is trying to individually find the best solution to a search problem. This provides a unique method for studying the

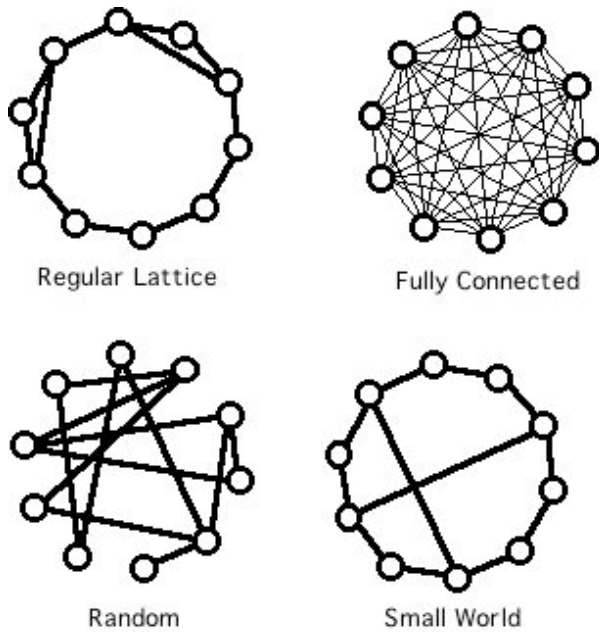


Figure 1: Examples of the different network structures for groups of 10 participants. Circles represent participants, lines indicate communication channels.

properties of networks using actual human behavior. In choosing a paradigm for studying information dissemination, we sought to find a case with: 1) a problem to solve with answers that varied continuously on a quantitative measure of quality, 2) a problem search space that was sufficiently large that no individual could cover it all in a reasonable amount of time, and 3) simple communications between participants that would be amenable to computational modeling. We settled upon a minimal search task in which participants guess numbers between 0 – 100 and the computer tells them how many points were obtained from the guess. There was a continuous function that related the guesses to the points earned, but this function was not revealed to the participants. The participants received information on their own guesses and earned points, as well as obtained information on their neighbors’ guesses and outcomes. In this manner, participants could choose to imitate high-scoring guesses from their peers. Examples of the network structures we compared are shown in Figure 1 for groups of 10 participants. Circles indicate participants and lines connect participants that directly exchange information. To generate our small-world networks we started with a spatially ordered network (i.e., the ring structure of the regular lattice) and added edges between random nodes. Although this caused the clustering coefficient to be low, because neighbors of a node were not more likely to be neighbors of each other, they still had a small average geodesic path length and maintained the spatial structuring of the lattice network, as evidenced by the greater closeness centrality of the networks.

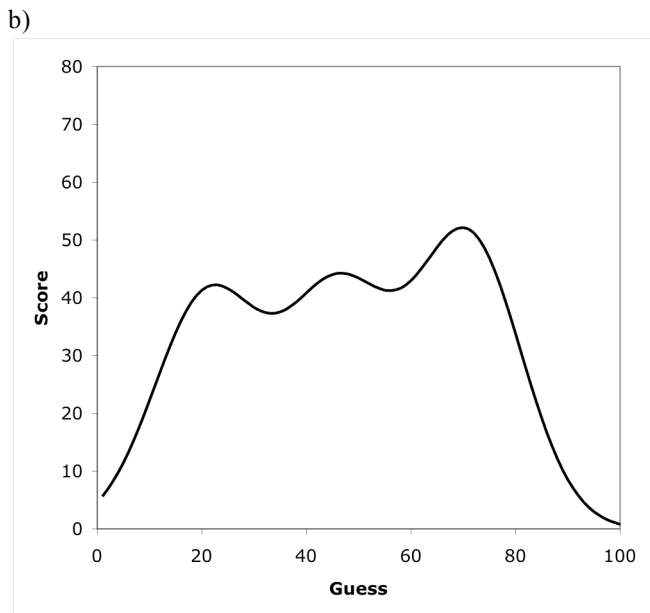
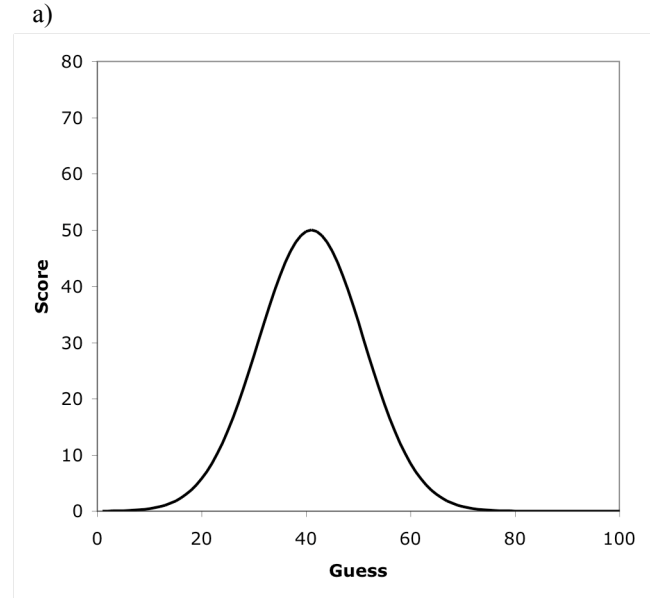


Figure 2: Examples of the a) unimodal and b) multimodal fitness functions

Notice in Figure 1 that three of the networks have a total of 12 connections between participants. Thus, if there is a difference in information dissemination in these networks, then it must be due to the topology, not density, of the connections. In addition to these three network structures, we also used a fully connected network (also called a “complete graph” in graph theory), in which everyone had access to the guesses and scores of everyone else.

In this experiment, we compared two fitness functions. The unimodal function has a single peak that could be found with a hill-climbing method. The multi-modal function increased the difficulty of the search and introduced local maxima. A local maximum is a solution that is better than

all of its immediate neighboring solutions, yet is not the best solution possible. In this case a simple hill-climbing method might not find the best possible solution. Figure 2 shows one of the multi-modal functions used, which has three peaks, but one of the peaks is somewhat higher than the other two. One interesting prediction is that participants may prematurely converge on a local maximum, thus precluding exploration of better, uninhabited regions of the problem space.

The basic prediction is that the small-world network will allow fast dissemination of information that will lead to rapid convergence on a maximum, but the early distribution will be divided into clusters that allows for more exploration and thus less likelihood of early convergence on a local maximum when compared to less cliquish network configurations like the fully connected networks. In the unimodal landscape there are no local maxima, so we expect that when the spread of information is fast (i.e., shorter average path length), convergence on the maximum will also be fast.

Method

112 groups of Indiana University undergraduate students ranging in size from 5 – 18 people with a median of 12 people per group participated for partial course credit, for a total of 1358 participants. Each session was run in a computer lab with 20 client computers used by the participants and one server operated by the experimenter. Before each session, the experimenter set up the server with one of two fixed random orderings of 8 series, each of which had a different fitness function and network structure. The fitness functions in each series were either unimodal or trimodal, but the positions of the global maximum (and local maxima in the trimodal conditions) were different for each of the 8 series. The network structure in each series was either full, lattice, small-world, or random, similar to those in Figure 1.

To create a network, the server takes all of the client computers and treats each as a node. For the random network, the server creates a number of edges equal to 1.3 times the number of nodes connecting random nodes under the constraint that a path exists between every node (i.e., that the graph is connected). This is conceptually equivalent to the algorithm proposed by Malloy and Reed (1995) for generating random networks with a pre-defined degree distribution. For the lattice network, the server connects the clients in a ring and then randomly picks 30% of the nodes and connects each of these nodes to a neighbor two steps away. For the small-world network, the server begins by putting the clients in a ring and then picks 30% of the nodes randomly and adds a connection to another random node under the restriction that the connected nodes are at least 3 nodes apart following the lattice path. These probabilities ensure that the average degree is roughly equivalent for all of the network structures. Thus, the experiment was a 2 (fitness function) x 4 (network structure) within-subjects (or rather, within-groups) design.

Participants signed onto the computer and gave themselves a handle or were assigned an ID. Once they had signed onto the computer, the experimenter started the session and the following instructions appeared to each of the participants:

Thank you for participating in this experiment on how ideas move from person to person in a social group. Your task is to try to accumulate as many points as possible. On each trial, you will type in a number between 0 and 100, and the computer will tell you how many points your number receives. There is a systematic relationship between the number you put in, and the points you receive, but the relationship will often be difficult for you to understand. Every time you type in the same number, it will be worth about the same number of points, but there may also be a bit of randomness added in to the earned points. Usually, numbers that are close to each other will receive similar points. At the end of each block of trials, you will be told how many points you earned, and how many points people earned in general.

In addition to telling you how many points your guess was worth, after each round of guesses, the computer will show you what numbers other people guessed, and how many points those guesses earned. You can use this information to help you decide what number to guess on the next round. Other people will also see the number that you entered, and how many points you received.

After they read this, the controlling program created the network neighborhood and the first round in the first series began. Each series consisted of 15 rounds in which participants had 20 seconds to guess a number between 0 - 100. When a round ended, the guesses were sent to the server, which calculated each participant's score (which was always between 0 and 50), added normally-distributed noise with a variance of 25, and returned the feedback. This began the next round, and participants now had available a list of the their own and their neighbors' ID, guess, and score while they decided on their next guess.

Results

We examined several measures of search performance to compare the different network structures on different fitness functions. To determine speed of convergence we looked at the average number of guesses the agents made before reaching the global maximum. To compare overall performance we looked at the percentage of the participants within one standard deviation of the global maximum on each round and across all rounds. To see how clustered the guesses were, we used the Kullback-Leibler¹, or relative entropy statistic, to compare the spread of guesses to a uniform distribution. The relative entropy increases as the distribution of guesses deviates more from a uniform

¹ The Kullback-Liebler is $\sum p_i \log(p_i/q_i)$ where p_i is the actual frequency and q_i is the expected frequency of guesses in each "bin" summed from $i = 0$ to N , or the number of bins that segment the range of guesses.

distribution. For our purposes we divided the range of guesses from 0 - 100 into 20 bins of 5 points each.

While the average number of rounds before any single individual reached the global maximum did not differ between network structures for the unimodal function, the lattice network had a significantly smaller percentage of people ($M = 0.66$) within the global maximum than all of the other networks (full: $M = 0.82$; small: $M = 0.80$; random: $M = 0.77$) ($F(3,50) = 4.191, p = 0.01$), which is most likely due to the longer average path lengths and thus slower spread of information through the network.

This is supported by examination of the percentage of participants within one standard deviation of the global maximum after each round. As can be seen in Figure 3a, the fully connected, small-world, and random networks have almost 80% of participants in the global maximum by Round 3, while the lattice network does not reach this level until round nine, and never reaches the percentages the other networks attain. Additionally, the low average relative entropy of the lattice network ($M = 1.43, SD = 0.42$) compared to all of the other networks (full: $M = 1.71, SD = 0.36$; small: $M = 1.61, SD = 0.45$; random: $M = 1.71, SD = 0.41$) indicates that the distribution of guesses in the lattice network was typically less clustered than for the other networks ($F(3, 791) = 17.259, p < 0.001$).

In the multimodal landscape we again expect shorter path lengths to correspond with faster convergence on the global maximum, but we anticipate that lack of spatial structure could lead to less exploration, and thus early convergence on a local maxima and a slower convergence on the global maximum.

As predicted, the average number of steps for the first person to reach the global maximum was significantly less in the small-world network ($M = 3.47$) than even the fully-connected network ($M = 5.10$), and this difference was significant ($t(21) = 2.9, p < 0.05$). The efficiency of the small-world network is also evidenced in the average percentage of participants within the global maximum across all rounds ($M = 0.674$), which is significantly higher than the other network structures (full: $M = 0.53$; lattice: $M = 0.43$; random: $M = 0.40$) ($F(3, 53) = 5.692, p < 0.005$). An examination of the percentage of participants within the global maximum on each round highlights the advantage of the small-world network. As can be seen in Figure 3b, the small-world network consistently dominates the other network structures, approaching the highest percentage (80%) by round nine while the other networks do not reach that level until the last round, if at all.

Discussion

When there was only one good solution – when the fitness function was unimodal – there was a direct relationship between the average shortest path length and the speed with which the group converged on the best solution. In this case, the fully connected network performed only slightly and nonsignificantly better than the random or small-world networks, as is predicted by the approximately equally short

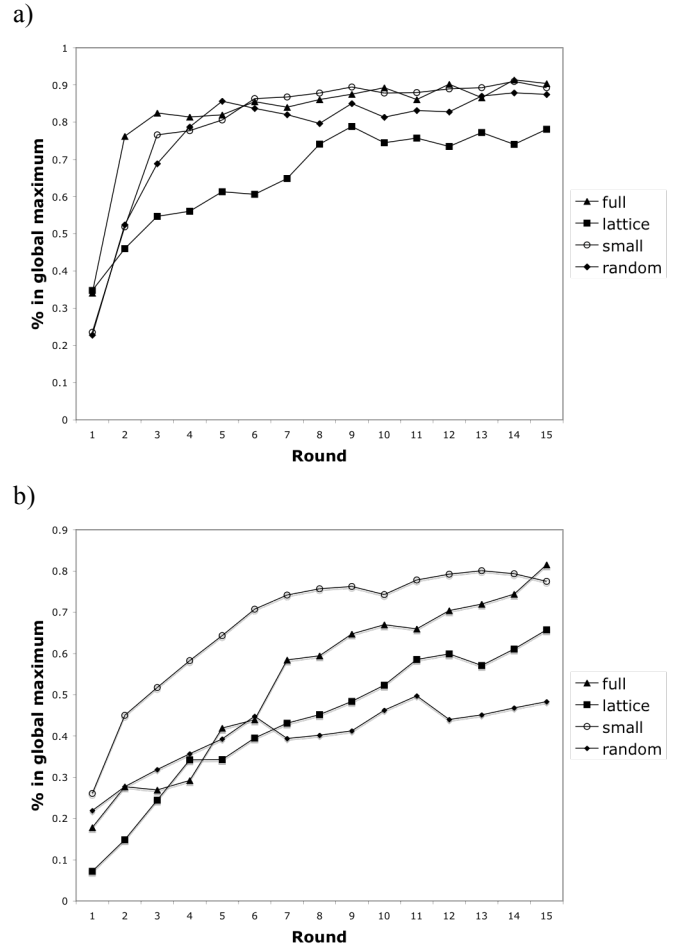


Figure 3: Percent of participants within 1 standard deviation of the global maximum on each round in the a) unimodal and b) trimodal fitness function.

path lengths for these three networks. The lattice network took longer to converge on the best solution because the advantageous innovations had to work their way through longer chains of people. However, when the problem space had good solutions that were nonetheless sub-optimal, as in the multimodal fitness function, the story was different. In this case the small-world network groups found the best solution faster than every other network, even the fully connected network in which everyone had the information about every other participants' guesses and scores.

This somewhat counter-intuitive result, that limiting the available information might actually improve a group's performance, is a result of the way the groups were searching the problem space. In the fully-connected network, participants would latch onto the first good solution that was found, and this was only the best solution a third of the time. When the group converged prematurely on a local maximum, it took them longer for an adventurous (or bored) participant to explore and find the globally best solution. In the small-world network, however, the participants were segregated into different spatial regions,

but the information could travel quickly through “short-cuts,” allowing for different locally connected groups to explore various regions of the problem space. Thus, while one locally connected group might latch onto a local maximum, the small-world topology decreases the probability that everyone will follow their lead before another sub-group finds the global maximum. Once any subgroup finds the global maximum, the information can spread quickly to other subgroups, unlike the lattice structure.

Ultimately, the paradigm developed here can be used to study the problem-solving abilities of groups under a wide range of conditions. For instance, different communication structures could be tested, such as scale-free networks (which are increasingly observed in a wide range of real networks), or hierarchies, which are interesting because they are a typical organizational structure. Additionally, different problem spaces remain to be explored, including multidimensional and dynamically evolving problem spaces. It seems reasonable to predict that a network structure that permits a group to quickly converge upon a solution may be less fit when the problem space changes.

Acknowledgments

This research was funded by NSF grant 0125287 to the third author.

References

- Albert, R., Jeong, H., & Barabasi, A.-L. (1999). The diameter of the World Wide Web. *Nature*, *401*, 130-131.
- Axelrod, R. (1997). The dissemination of culture: A model with local convergence and global polarization. *Journal of Conflict Resolution*, *4*, 203-226.
- Bandura A. (1965). Behavioral modification through modeling procedures. In L. Krasner & L. P. Ulmann (Eds.), *Research in behavior modification: new development and implications*. (pp. 310-340). New York : Rinehart and Winston.
- Banerjee, A. V. (1992). A simple model of herd behavior. *Quarterly Journal of Economics*, *107*(3), 797-817.
- Bikhchandani, S., Hirshleifer, D., & Welch, I. (1992). A theory of fads, fashions, custom, and cultural change as informational cascades. *The Journal of Political Economy*, *100*(5), 992-1026.
- Blackmore, S. J. (1999). *The meme machine*. Oxford: Oxford University Press.
- Chwe, M. S-Y. (1999). Structure and strategy in collective action. *American Journal of Sociology*, *105*, 128-156.
- Cialdini, R. B. & Goldstein, N. J. (2004). Social influence: Compliance and conformity. *Annual Review of Psychology*, *55*, 591-621.
- Deutsch, M. & Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgment. *Journal of Abnormal and Social Psychology*, *51*, 629-636.
- Erdős, P. & Rényi, A. (1959). On random graphs. *Publicationes Mathematicae*, *6*, 290-297.
- Freeman, L. C. (1979). Centrality in social networks: Conceptual clarification. *Social Networks*, *1*, 215-239.
- Gladwell, M. (2000). *The tipping point*. London: Little Brown and Company.
- Granovetter, M. (1978). Threshold models of collective behavior. *American Journal of Sociology*, *83*, 1420-1443.
- Holland, J. H. (1975). *Adaptation in Natural and Artificial Systems*. Ann Arbor: University of Michigan Press.
- Jeong, H., Tombor, B., Albert, R., Oltvai, Z. N., & Barabasi, A.-L. (2000). The large-scale organization of metabolic networks. *Nature*, *407*, 651-654.
- Kennedy, J., Eberhart, R. C., & Shi, Y. (2001). *Swarm intelligence*. San Francisco: Morgan Kaufmann Publishers.
- Kleinberg, J. (2000). Navigation in a small world. *Nature*, *406*, 845.
- Malloy, M. & Reed, B. (1995). A critical point for random graphs with a given degree sequence. *Random Structures & Algorithms*, *6*, 161-179.
- Milgram, S. (1967). The small world problem. *Psychology Today*, *2*, 60-67.
- Newman, M. E. J. (2001). The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences*, *98*, 404-409.
- Nowak, A., Szamrej, J., & Latané, B. (1990). From private attitude to public opinion: A dynamic theory of social impact. *Psychological Review*, *97*, 362-376.
- Rogers, E. M. (1962). *Diffusion of Innovations*. New York: Free Press.
- Rogers, E. M. (1995). *Diffusion of Innovations (4th ed.)*. New York: Free Press.
- Ryan, B. & Gross, N. C. (1943). The diffusion of hybrid seed corn in two Iowa communities. *Rural Sociology*, *8*, 15-24.
- Sherif, M. (1935). A study of some social factors in perception. *Archives of Psychology*, *187*, 60.
- Valente, T. W. (1996). Social network thresholds in the diffusion of innovations. *Social Networks*, *18*, 69-89.
- Watts, D. J. & Strogatz, S. H. (1998). Collective dynamics of “small-world” networks. *Nature*, *393*, 440-442.
- Wilhite, A. (2001). Bilateral trade and ‘small-world’ networks. *Computational Economics*, *18*, 49-64.