Siddharth Suri
www.sidsuri.com
sidsuri@ymail.com

Research Statement

The Internet, the World Wide Web, Facebook, and eBay are examples of some of the myriad types of networks that are a part of everyday life for many people. Three important types of networks are technological, social, and economic. First, an example of a technological network can be seen in a model of the World Wide Web where a node could represent a web page and edges could represent hyperlinks. Similarly, in a model of the Internet nodes could represent routers and edges could represent the exchange of traffic. Second, social networks encompass a broad range of networks that occur in the real world where the nodes represent some type of naturally occurring entity and links between pairs of nodes represent some type of social relationship. For example, there are social networks where nodes can represent people and links can represent a friendship, a past communication, a coauthorship, or a citation. Third, economic networks are similar except they use links to model some type of economic interaction between nodes. Examples of economic networks include models where each node is a person, a firm, or a country and links represent trade agreements.

Since these networks are often inherently technological and involve social and/or economic interactions between the nodes, it is necessary to use theories and models from three fields — sociology, economics, and especially computer science — to study them. Theories from sociology are needed to understand the relationships that people form and maintain over these networks. Models from economics are needed to understand the interactions of agents who act in their own self-interest. Computer science, in particular, could have a broad impact on studies of social networks as well as economic networks. For example, analyzing dynamics over networks and how they affect, and are affected by, the topology of the network provides one broad class of computational questions. A more specific algorithmic question in this general area is to study the influence people exert on each other to adopt a new behavior and how the spread of the new behavior is affected by the topology of the network. A second broad class of computational questions is motivated by networks where the interactions are well described by a game theoretic model. A natural computational question is to analyze the quality of equilibrium when compared to a socially optimal solution. Another example in this general category of questions is quantifying how the topology of the network affects the payoff of the players. Understanding the various types of interactions and dynamics over networks is key to increasing scientists' understanding of how to build and maintain more efficient and effective systems. My research contributes to this endeavor by focusing on the algorithmic and computational issues involved in the study of social networks and economic networks.

Past Research My thesis work combined structural models of social networks with models of interaction from economics and analyzed this combination from a computational perspective. The field of social network theory[1] primarily focuses on measuring the topology of naturally

[1] Since this field is relatively new, its name has not yet been codified. I use the term social network theory to refer to the study of interactions between members of a large population; it has alternately been referred to as *network science* [3]. This area represents the confluence of computer science, sociology, economics, applied mathematics, physics, and related fields.

occurring networks and then inventing generative models that output networks of similar structure (see for example [13, 14]). These generative models are often purely structural. That is, they model patterns of links between nodes in networks rather than attempting to model the behavior of the people that these nodes represent. On the other hand, the field of economics and, specifically game theory, describes many models of interaction. Game theory analyzes models of fully rational agents acting in their own best interests (see for example [12, 2]). The overwhelming majority of these models, however, do not consider network structure. They assume that all players can interact with and directly affect all other players. Combining the interaction models of game theory with the network models of social network theory yielded a much richer, more accurate model of human behavior. My thesis work began the exploration of this broad new area by addressing one fundamental question: if players are arranged in a network, and they are strategically interacting only with other players in their local neighborhood, how does the topology of the network affect the outcome of the interaction? My thesis answered this question in a variety of different contexts necessitating the use of a variety of analytical techniques.

First, my dissertation research analyzed a graphical model of markets. This model consisted of a graph where trade could only occur between two parties if they were attached by an edge. We quantified the variation of wealth at equilibrium in social networks in terms of the degree distribution of the underlying network. It turned out that a player could achieve a high equilibrium wealth solely due to its position in the the network. Moreover, my coauthors and I used simulations to test how tight the bounds in our theorems were. We also used experiments on real world international trade data to support and motivate our model. Subsequently, we theoretically analyzed an extension of this model where the players must first buy the edges in the network and then trade according to the network that was purchased. We gave a complete characterization of the equilibrium networks in this model. These results, which were originally published in [7] and [5] respectively, fundamentally related the network topology to the equilibrium behavior.

Second, we assessed a model of evolutionary game theory over networks. Evolutionary game theory has been used to model biological and social interactions where agents are not assumed to be fully rational. Instead, agents are simply assumed to copy the behavior of an agent in their local neighborhood who is doing better then they are. For two broad classes of graphs, we completely characterized which strategies could be played by a large fraction of the population that would guarantee they could not be overrun by any small mutant invasion. The topology of the networks under study was paramount in characterizing which strategies are resilient to invasion. This theoretical work, which first appeared in [8], integrates game theory, as a model for social interaction, with graph theory.

The final component of my dissertation research was more experimental in nature. Its purpose was to investigate how the same group of people behaved when organized in different networks. A group of 38 undergraduates were instructed to solve the graph coloring problem. Each one used our asynchronous software system to control the color of one node in a network. Their objective was to arrive at a valid coloring (where each node has a color different than each of its neighbors) in under 300 seconds. The subjects were paid for their performance. We systematically varied the topology of the networks that the subjects were given to color. They colored networks chosen from two generative models of social networks. One of the key findings of this experiment was that the group could color networks with low diameter much faster than networks with high diameter — again showing that network topology impacted group behavior. This work involved experiments on actual human subjects and statistical analyses of the data gathered from these experiments. It originally appeared in [9].

As a postdoctoral researcher at Cornell University I extended this line of study using large scale data analysis. My coauthors and I analyzed how a person's local neighborhood influences their behavior. More specifically, we compared the probability that a person will adopt a behavior given that k of his/her neighbors have already adopted that behavior to the probability that a person will adopt a behavior given that k of the people most similar to that person have already adopted the behavior. This work focuses on measuring the influence caused by the people that an individual communicates with in relation to the behavior of those people in the population who are similar to that individual. This research [4] involved the analysis of the edit history of Wikipedia and crawls from the LiveJournal online community.

My work at Cornell also included a project that was more theoretical in nature. Sociologists have discovered that people can gain from bridging two separate communities [1]. For example, a scientist who bridges two fields can apply ideas from one field to solve open problems in the other. That scientist could also innovate by combining ideas from two disparate fields into something new. Thus, a person's position in a network can directly impact their livelihood. Studies of this phenomenon typically analyze existing, static networks. We modeled how a network with this topological feature might arise organically using a network formation game. In [10] we encoded the desire to bridge two communities in the players' utility functions and analyzed what types of networks form at equilibrium.

Future Research My future work will primarily focus on understanding how the behavior of individuals affects the overall topology of the network. In other words, I will study how the global structure of the network emerges from the local dynamics between the nodes. Many empirical studies have shown that networks from a wide variety of contexts (e.g. sociology, biology, technology, economics, etc.) have similar statistical properties such as a heavy tailed degree distribution, an abundance of triangles, small diameter, along with other related properties. I intend to take a bottom up approach to understanding how these properties arise by arriving at accurate models of how nodes attach to other nodes and how nodes interact with other nodes. These interactions may depend on the context of the networks in question, thus different models may be necessary for different classes of networks. As discussed above, previous work in the social network theory literature has derived stochastic models that output networks that share statistical properties with real-world networks. But almost all of these models fail to consider the interactions occurring between nodes. The literature on network formation games (see [6] for an excellent survey) on the other hand, assumes that nodes are fully rational players who choose which other nodes to attach to according to their individual utility functions. The networks that arise at equilibrium have typically a very rigid and symmetric structure which is not reflected in topologies seen empirically in the real world. My future work intends to resolve this dichotomy in our current understanding of social networks.

My approach to filling in this gap in the literature will utilize a diversity of techniques including: analysis of large data sets, behavioral experiments, theoretical analysis, and simulation. In order to arrive at accurate models of interaction in social networks we must first understand the behavior of the agents, whether they be people, organisms, machines or other entities. The rise of the Internet has made available large sets of data that allow one to study the dynamics of the nodes in a variety of types of networks. I intend to use large scale data analysis, similar to the analysis I have done on Wikipedia and LiveJournal and am currently doing on Flickr, to understand these dynamics. Analyzing the data from these sources will unearth various computational questions. For example, this type of analysis will require new algorithms that do more than simply analyze massive graphs, they will also need to analyze data from the interactions that occur over these graphs. In addition, techniques from machine learning will be necessary to abstract models from these data.

I also intend to use more controlled behavioral experiments to improve our understanding of the specific dynamics between pairs of people arranged in a network. Typically, laboratory controlled behavioral experiments yield much less data than can be gathered from online sources, but the data is relatively low in noise and is more likely to confirm or refute a specific hypothesis. The field of behavioral game theory yielded a paradigm for conducting such experiments, which I have successfully used in the past and will continue to use in the future. As an example of this, in my current resarch at Yahoo!, I am using human subject experimentation to study how network topology influences contribution levels in public goods games.

To complement these approaches, I intend to use both theoretical analysis and simulations in my study of social and economic models of interaction over networks. A theoretical study of these types of models would allow us to characterize the different types of interactions that occur in networks of different contexts. This would allow for the study of general classes of social networks as opposed to inventing new theories for each individual social network. Furthermore, formally analyzing the local dynamics would lead to a better understanding of how changes to local incentives and local interactions effect the behavior of a population as a whole, and in turn, the general topology of the network.

Simulations can be used to exhibit the dynamics of the agents in these networks. For example large scale simulations could be used to incorporate the behavioral models into a large network setting and then check that the macro-level phenomenon observed by the large scale data analyses emerge. This would provide a validation of the more focused behavioral models as well as exhibit how they combine on a large scale. Moreover, simulations are also useful for showing how to control and manipulate the population [11]. Thus, simulations are interesting in their own right and they can also be used to elucidate equilibrium behavior, which is helpful in the formal analysis of game theoretic models.

In summation, the intersection of social network theory and economics, when studied from a computer science point of view, provides a timely new research area teeming with unexplored, well motivated research directions. It is only recently, with the proliferation Internet usage, that it has been feasible to gather large data sets to inform this study. It is also only very recently that scientists have had the computational power readily available to make study of data sets of these magnitudes feasible and to perform simulations on a scale large enough to understand the dynamics of these networks. Furthermore, the relatively young field of behavioral economics has gained traction and provided us with the protocols for doing experiments on actual human subjects. Finally, theoretical techniques for the study of models of network behavior have been developed in the emerging field of algorithmic game theory. Thus the intersection of computer science, social network theory, and economics is currently ripe for exploration, and my research agenda sits squarely at their nexus.

References

[1] Ronald S. Burt. Structural holes and good ideas. *American Journal of Sociology*, 110(2):349–399, September 2004.

[2] Colin F. Camerer. *Behavioral Game Theory*. Princeton University Press, 2003.

[3] Committee on Network Science for Future Army Applications. *Network Science*. The National Academies Press, 2005. National Research Council.

[4] David Crandall, Dan Cosley, Daniel Huttenlocher, Jon Kleinberg, and Siddharth Suri. Feedback effects between similarity and social influence in online communities. In *Proceedings of the 14th Annual ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2008.

[5] Eyal Even-Dar, Michael Kearns, and Siddharth Suri. A network formation game for bipartite exchange economies. In *Proceedings of the 18th Annual ACM-SIAM Symposium on Discrete Algorithms*, 2007.

[6] Matthew Jackson. A survey of models of network formation: Stability and efficiency. In Gabrielle Demange and Myrna Wooders, editors, *Group Formation in Economics; Networks, Clubs and Coalitions*. Cambridge University Press, 2005.

[7] Sham M. Kakade, Michael Kearns, Luis E. Ortiz, Robin Pemantle, and Siddharth Suri. Economic properties of social networks. In *Proceedings of the Eighteenth Annual Conference on Neural Information Processing Systems*, 2004.

[8] Michael Kearns and Siddharth Suri. Networks preserving evolutionary equilibria and the power of randomization. In *Proceedings of the 7th ACM Conference on Electronic Commerce*, 2006.

[9] Michael Kearns, Siddharth Suri, and Nick Montfort. An experimental study of the coloring problem on human subject networks. *Science*, 313(5788):824–827, August 2006.

[10] Jon Kleinberg, Siddharth Suri, Éva Tardos, and Tom Wexler. Strategic network formation with structural holes. In *Proceedings of the 9th ACM Conference on Electronic Commerce*, 2008.

[11] Michael Mitzenmacher. The future of power law research. *Internet Mathematics*, 2(4), 2005-2006.

[12] Martin J. Osborne and Ariel Rubinstein. *A Course in Game Theory*. The MIT Press, 1994.

[13] Duncan J. Watts. *Small Worlds: The Dynamics of Networks between Order and Randomness*. Princeton University Press, 1999.

[14] Duncan J. Watts. *Six Degrees: The Science of a Connected Age*. W. W. Norton & Company, 2003.