Do Social Networks Improve e-Commerce? A Study on Social Marketplaces

Gayatri Swamynathan, Christo Wilson, Bryce Boe, Kevin Almeroth, and Ben Y. Zhao
Department of Computer Science, UC Santa Barbara, CA 93106
{gayatri, bowlin, bboe, almeroth, ravenben}@cs.ucsb.edu

ABSTRACT

Social networks have made a significant impact on how Internet users communicate, search for and share data today. Numerous proposals have been made to improve existing distributed systems by leveraging the inherent trust built into social links. For example, many believe that by augmenting online marketplaces with social networking, we can improve trust between transaction partners and increase user satisfaction. In this paper, we perform a detailed study of Overstock Auctions, an auction site that has recently integrated social links into user profiles. Using data on connections between roughly 400,000 Overstock users, we evaluate the impact of social connections on business transactions. Our results show that while the majority of users do not engage in social networking, those who transact with friends from their social network generally obtain significant benefits in the form of higher user satisfaction.

Categories and Subject Descriptors

J.m [**Computer Applications**]: Miscellaneous

General Terms

Human Factors, Measurement, Performance

Keywords

e-Commerce, Social Networks, Trust

1. INTRODUCTION

Social networks such as MySpace, Facebook, and YouTube have made a significant impact on how today's Internet users communicate, search for and share data. Users join these networks, publish and maintain their own profiles, and establish links to their friends. The resulting social links are evidence that a level of trust has been established between the connected users.

Numerous proposals have been made to improve existing distributed systems by leveraging the inherent trust built into social links. In particular, many believe that augmenting online marketplaces with social networking features should improve trust between transaction partners and improve user satisfaction. Customers on traditional online marketplaces such as eBay and Craigslist routinely transact with complete strangers, thus making them vulnerable to malicious cheaters out to defraud them. Despite their integration of reputation systems, prior work shows malicious users can artificially inflate their reputation values [3]. In contrast, a "social marketplace" would integrate features of social networks into an online shopping community, allowing customers to seek out purchases from their friends or friends of friends (FOF).

To evaluate the potential impact of social networking on online marketplaces, we study user connectivity and behavior through a large dataset of users from Overstock Auctions. Overstock Auctions is an online auction house similar to eBay and uBid, but is unique in its integration of a social network into the market community. Like other social networks, Overstock Auctions encourages users to establish an online presence through personalized homepage with personal history, photos and links to friends. Unlike typical social networks, Overstock Auctions user profiles often include their shopping preferences and return policies.

Overstock Auction maintains two distinct networks of its members. Users can become members of the social network by becoming friends with other users. Users who buy or sell items become part of the business network, where each link connecting two users implies one or more prior financial transactions. To guide their financial decision-making, the website offers users information on how they are connected to each potential transaction partner.

In this paper, we describe results from our analysis of a large anonymized user dataset obtained from Overstock Auctions. The dataset includes over 400,000 Overstock users and their peer-to-peer links in both the social and business networks. We study several aspects of Overstock's social network, including its level of adoption, the connectivity of business partners on the social network, and differences in user satisfaction following transactions between users connected via social links. We find that while the majority of users do not engage in social networking, those who transact with friends of friends generally obtain significantly benefits in the form of higher user satisfaction. Our results show that, with sufficient adoption, social networking can have a dramatic positive impact on online marketplaces.

The rest of the paper is structured as follows. We provide background information about Overstock Auctions in Section 2, followed by our data crawling methodology in Section 3. We then describe our key findings and results in Section 4. Finally, we discuss related work in Section 5, and conclude in Section 6.

2. CASE STUDY: OVERSTOCK AUCTIONS

In this section, we provide background information on Overstock Auctions, its social networking component. and the reputation system it uses to maintain a web of trust between participating users.

WOSN'08, August 18, 2008, Seattle, Washington, USA.

Table 1: Social network data provided by Overstock

	Business Network	Social Network
Total network nodes	398989	85200
Total network links	1926553	1895100
Avg node degree	4.82	22.24
Link symmetry	Yes	Yes

Table 2: Crawled transaction logs format

SellerID	BuyerID	TID	Date	Rating
A	B	123	2/19/2005	+2
A	B	234	12/17/2004	+2
A	C	345	12/15/2004	0
B	D	456	12/2/2004	-1
C	F	678	12/1/2004	+1

Overstock is an online retailer that offers brand-name merchandise at discounted prices over the Internet. *Overstock Auctions*, the online auction platform of Overstock, launched in 2004, supports a large community of users worldwide.

Similar to eBay's Feedback Forum, buyers and sellers on Overstock Auctions[1] rate one another at the end of each transaction, and these ratings are aggregated to form a user's feedback-based reputation profile. Additionally, users of Overstock Auctions are invited to create free personal pages, post photos, list hobbies and interests, and publish friends and business contact lists. This social networking component is designed to build a community across buyers and sellers, thereby improving trust and improving transaction satisfaction beyond what is possible with a reputation system.

Overstock's reputation system. Feedback ratings (-2 to +2) are aggregated by Overstock to form a user's "business rating" and "positive percentage" score. The business rating is the sum of the average rating received from each distinct transaction partner. The positive ratings percentage reflects the percentage of non-negative transactions in which a user has been involved. Overstock also provides transaction summaries for each user such as the total number of transactions undertaken by the user, the percent of items bought by a repeat customer, etc.

Overstock social networks. A user's "personal network" comprises of friends that have been invited by the user to be part of its network. Friends can choose to rate each other on a 1 to 5 scale. Overstock provides the ability for each user to view its web of trust leading to a particular buyer or seller. Users who trust the opinions of friends, family, or other social entities can choose to rely upon the personal friends-of-friends (FOF) network to help make connections and guide decisions.

Overstock provides another category of social network called a "business network." Every time an Overstock user completes a transaction, the transaction partner becomes a part of its first degree business network, thereby, creating a viral marketing effect as each new business partner is in turn connected to its business network.

3. MEASUREMENT METHODOLOGY

Our analysis relies on two datasets. The first is a large dataset provided by Overstock consisting of the personal and business network connectivity data for 431705 users. This includes the set of all currently "active" users, but does not include members who have not made transactions or users who are no longer active. Table 1 summarizes the network graph statistics for the business and personal networks provided by Overstock. The remaining dataset is a set of 10,000 user profiles downloaded from Overstock. The remainder of this section describes our data gathering methodology.

The Overstock Crawler The large Overstock dataset allows us to empirically examine the impact of business and personal networks on the operation of Overstock's online marketplace. However, we required specific feedback histories of users to evaluate user satisfaction after transactions. We downloaded detailed transaction histories of 10000 Overstock users, including their individual transaction feedback ratings, aggregated business rating and positive feedback percentage.

The Overstock crawler is a Python application structured around a core multi-threaded engine that handles all low-level socket and HTTP transactions with the target website. We implement a modular "handler" class that is subclassed to implement specific functionality for parsing pages from a website like Overstock. Additionally, we implement an "authorization module" integrated with the core engine to enable support for website-specific user authentication schemes. The crawler is fully distributed and employs the Python RMI to centralize the queue of uncrawled page URLs.

Starting with an initial seed of 15 randomly selected Overstock users, we obtain transaction histories of these users, and add their personal and business contacts to our breadth-first search. Overall, our crawler downloaded over 75000 webpages, parsing data that represented detailed transaction histories of over 10000 individual Overstock users. To avoid having a negative effect on regular Overstock users, we rate limited our download traffic to 1 Mbit/second.

We list the format of Overstock transaction logs in Table 2. For each transaction, we log the buyer and seller identities, transaction id (as given by *TID*), transaction date, and finally, the rating (-2 through $+2$) given by the buyer to the seller for the given transaction. Furthermore, we extract the overall feedback summaries presented on each user's feedback profile page, namely, aggregate business rating, positive feedback percent, etc.

4. IMPACT OF SOCIAL NETWORKS

We now discuss detailed analysis of our data on Overstock. We are primarily interested in questions regarding the adoption of Overstock's social network by its users, correlation between number of transactions and social network distance, and correlation between per-transaction satisfaction and social network distance.

4.1 Overstock Connectivity Graph Analysis

We begin by characterizing the large Overstock connectivity dataset. The business network graph comprises 398989 users, while personal network graph includes 85200 users. We determined that 52484 users had both business and personal networks, implying that 86% of Overstock's business users did not possess a personal network at all. This shows that the large majority of users are primarily interested in financial transactions, and are either unaware or uninterested in the social networking component of Overstock. Given the value of social connectivity in ensuring successful transactions (discussed in Section 4.5), we believe Overstock can become more successful by further advertising the benefits of participation in its social network.

Interestingly, we also find that 38% users (~ 32716 users) had only a social presence on Overstock but have performed no business transactions. We believe the presence of these users are likely due to two factors. First, Overstock has an active message board

[1] We hereby refer to "Overstock Auctions" as "Overstock."

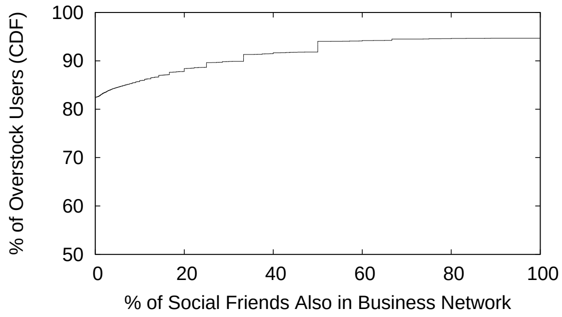

Figure 1: Number of friends common to a users's personal and business networks expressed as a percent of the user's personal network size.

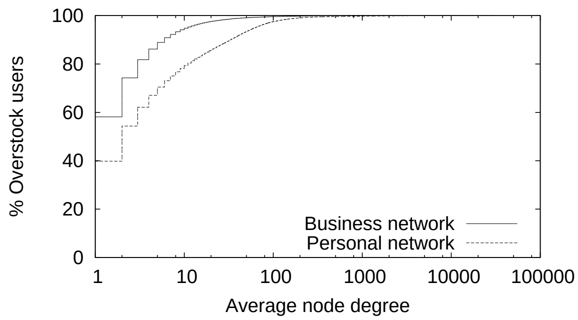

Figure 2: Node degrees for Overstock business and social networks.

system for discussions on any topics related to merchandise or the site in general. It is possible a number of users are interested in information from message boards, but have yet to make their first purchase. Second, social psychologists have recognized a phenomenon called *public displays of connections,* where social networking users use displayed connections as an indicator of popularity, and therefore form social links for the purpose of enhancing the popularity status of both parties [5]. In Overstock, social friends can rate each other. Thus having a larger group of social friends can boost the observed rating, thereby conveying an increased sense of trustworthiness and attracting more sales.

4.2 Overlap of Personal and Business Networks

We conduct a simple experiment to investigate whether Overstock users tend to have friends common to both their business as well as personal networks. Our goal is to observe the extent of overlap between these two networks. For each user, we compute the percentage of the users' social friends who are also in the user's business network, and plot the result as a CDF in Figure 1. As observed earlier, 86% users in the dataset have no personal network in place. For the remaining 85200 users, we observe that about 94% of users have less than 50% friends from their personal network overlapping with the business network. In fact, 82% users find less than 1% overlap of personal and business networks.

When we compute the inverse plot, however, we observe that 6% of users have 100% of their personal contacts participating in their business networks, indicating that these users have either transacted with all of their social friends, or invited business partners to join their personal network of friends. We omit the figure for space considerations.

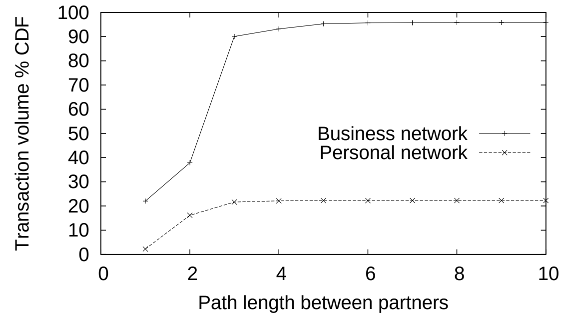

Figure 3: Relationship between percentage of transactions conducted and path length between partners in Overstock networks.

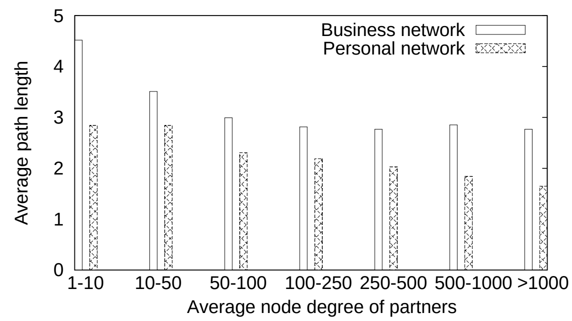

Figure 4: Relationship between average node degree and path length between two users in Overstock networks.

4.3 Exponential Degree Distributions

Figure 2 plots the average Overstock user's social network degree. We define the *degree* of a node in a social network graph as the number of links incident to the node, *i.e.*, a user's immediate personal and business friends in the personal and business network graphs respectively. An exponential degree distribution was observed for all networks. In addition, we derive the following statistics: the average node degree in the business network is 5 users; 99.55% of business network nodes (or users) have a degree less than 100, while roughly 92% nodes have degree less than 10. This shows that a significant portion of financial transactions are driven by a relatively small portion of Overstock users, who likely represent companies with storefronts in Overstock.

On the other hand, the average node degree in the personal network graph was 22 users, which suggests that users with personal networks are well connected within their social network. 99.7% of users have degree less than 1000, while 75% of users have degrees less than 10. The highest node degrees observed in the business and personal networks were in the range of 12000 and 30000 respectively, indicating that these members are power sellers or active social members of the Overstock community who, as our experiments later indicate, find themselves connected to a large fraction of the graph within a path length of 2 hops.

4.4 Average Path Lengths

In the following set of experiments, we determine the average path lengths between transaction partners on Overstock. We define path length as the length of the shortest path between two nodes in a network graph. "Personal path length" and "business path length" respectively refer to the shortest paths between two

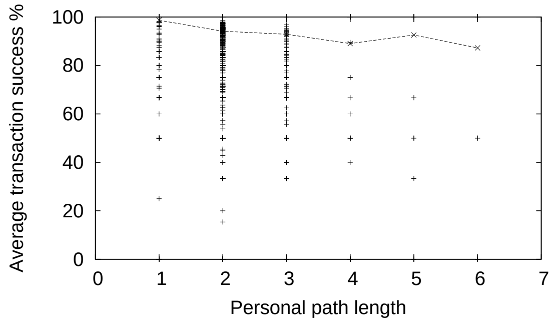

Figure 5: Correlation between transaction success and partner's social distance on the social network.

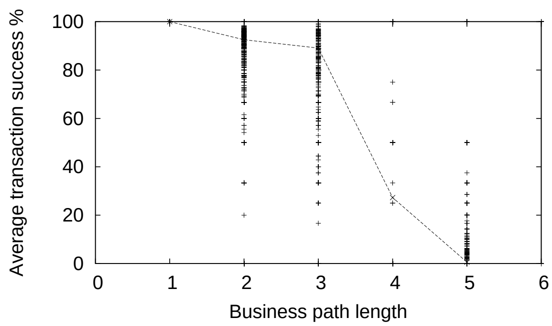

Figure 6: Correlation between transaction success and partner's distance on the business network.

users in the personal and business network graphs. We study the relationship of path lengths between transaction partners (both in personal and business networks) and the average volume of transactions observed for those path lengths. For instance, a high volume of transactions observed for short social path lengths could possibly indicate greater reliance of the Overstock community on friends-of-friends connections. We also analyze the relationship between node degrees and path lengths in the networks.

4.4.1 Computing Shortest Paths

Computation of shortest paths between transaction partners in a 400,000 node network is both memory and computationally expensive. To compute the shortest path length between transaction partners, we use an approximation heuristic by constructing a multi-level shortest-path friends-of-friends (FOF) tree for both nodes. These FOF trees produce shortest paths between the tree root and its leaves. Therefore, if node A resides in the FOF tree of node B, the shortest path length is the level of A in B's tree.

For each level of the FOF trees, we perform a set intersection of leaf nodes to locate common leaves. Any common node proves a path between the nodes equal to the sum of their tree depths minus one. We repeatly expand the depth of the FOF trees for each node until a common node is found between the leaves of the two trees.

We repeat the algorithm until we either detect a path between the buyer and seller, or until a maximum depth has been reached on the FOF trees. We set the FOF tree max depth to 6 in our experiments, thereby ignoring possible paths of length 11 or more. Our intuition is that path lengths greater than 10 hops in personal or business networks are minimally relevant to our studies of transaction success.

4.4.2 Correlation with Transaction Volume

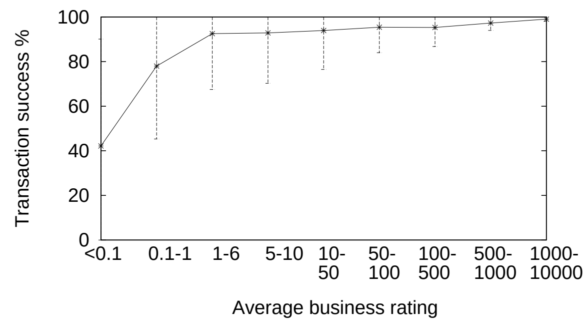

Figure 7: Correlation between Overstock's rating for a business and success rate of transactions conducted by the business.

Our first experiment determines average path lengths for each transaction pair found in our crawled transaction logs (Table 2). Since our crawl was completed in less than 6 hours, we assume that the network graph structures are static and did not grow or shrink over the course of our experiment, thereby enabling us to determine the impact of path lengths on each transaction. As part of future work, we propose to study the evolution of Overstock's graph structures over the course of network transactions.

A total of 17376 transactions were studied. For each path length observed between partners, we compute the "transaction volume percentage," *i.e.*, the total number of transactions observed for the given path length as compared to the total transactions. The following observations were made on Overstock personal networks. We observe that nearly 78% of transactions are between partners with path lengths greater than 11 hops, or one or both partners possessing no personal network at all($\sim$ 9% of this set). Clearly, these transaction partners are not relying on personal network connectivity to guide their decision to transact. As seen in Figure 3, approximately 2% of all business transactions were first degree personal friends. The users possessing a social network primarily transact with 2 to 3 hop partners (as observed from the $\sim$ 20% of remaining transactions) which we assume is because Overstock displays upto a 3-level FOF tree for transacting partners.

We next observe the average business path lengths. Because the logs depict transactions already undertaken in the network, transaction partners will already find themselves in each other's first-degree business network. Therefore, in this experiment, partners that have transacted together more than once assume a path length of 1 connecting each other. Otherwise, we remove them from each other's business network and re-compute the path lengths between them.

As observed with personal networks, the volume of transactions undertaken reduces as the path length between a transaction partner pair increases. However, we observe that users prefer business FOF relations significantly more than personal FOF relations. Furthermore, 22% partner pairs observe a business path of 1, thereby, indicating that users tend to prefer transacting with partners with whom they have interacted in the past. Only 4.2% transactions were observed to be carried between partners 11 hops away, the bulk of transactions being carried out between 2 to 5 hop business neighbors.

4.4.3 Relation with Node Degrees

We next study the relationship of the average node degree of transaction partners and the path lengths observed for those partners. We observe, in general, that the larger a pair's node degree, the smaller is the average path length connecting them (Figure 4).

Interestingly, a user can reach its partners in less than 3 hops using personal FOF even if has less than 10 first degree friends, while on the business network users would need 50 friends to be connected to partners 3 hops away. In order to trust business FOFs relations, users need to transact a lot more in the network to develop shorter paths and greater reliability to new transaction partners.

4.5 Transaction Success Rate

The objective of the following experiments is to empirically determine the impact of social networks on the transaction satisfaction experienced by buyers in the Overstock network. We deem a transaction "successful" if the buyer or seller has received positive ratings for the transaction. We define "transaction success rate" for each user as the percentage of successful transactions over the total transactions undertaken by the user. We examine the correlation between path lengths between transaction partners (both in personal and business networks) and the average transaction success percentages observed by the buyers. For both social and business networks, we aggregate all transactions based on hop distances separating the partners, and plot their average success rate as Figure 5 and Figure 6 respectively.

We again study the 17376 transactions generated from our crawl logs. Our previous experiment demonstrated that only 22% of all transactions are conducted between partners connected through the social network. While this ratio is low, we do observe that these partners generally find high success rates in their transactions. Partners less than 3 hops away from one another on the personal network graph find, on average, a 90% transaction satisfaction rate, indicating a high degree of success while relying on personal FOFs. In fact, even partners separated by 4 to 6 social network hops find an 80% or greater transaction success rate. We believe that connections between socially linked transaction partners serves to filter out unscrupulous cheaters in the system, since these cheaters are unlikely to have strong links to well-behaved users.

Interestingly, the same success does not occur for partners connected on the Overstock business network. Users do find high transaction success for business associates 3 or less hops away on the business network, but transaction success drops rapidly past 3 hops. This sharp drop in transaction satisfaction indicates that trust in the FOF business network degrades much faster than trust in the social network. While users can selectively choose other users to include in their personal networks, business FOF relations are only indicative of one or more prior transaction between the two users.

4.6 Accuracy of Overstock Business Ratings

We now investigate the benefits of traditional reputations as compared to social networks on Overstock. The objective of our first experiment is to compare transaction success experienced by the community with respect to Overstock's user "business rating" scores. Once again, we define "transaction success" for each user as the percentage of successful transactions (given by positive ratings) over the total transactions undertaken by the user. A user's business rating is the sum of the average feedback ratings (-2 to +2) received from distinct transaction partners. Business ratings can be arbitrarily positive or negative.

We aggregate all transactions observed in our dataset by the business rating of the seller. By grouping these sellers into groups, we compute the average transaction success rate per group, and plot the results in Figure 7. There is a clear and definitive trend in our plot, showing a strong correlation between business ratings and resulting transaction success rates. We first observe that business ratings less than 10 are generally unreliable: transaction success percentages vary from very poor to very high for these ratings. However, transactions with partners with higher business ratings (> 10) results in an increase in transaction success experienced by the community. Only 5% users experience transaction success percentages lower than 80% while transacting with users having > 100 rating score. Transacting with users having > 1000 aggregate business ratings score results in 90% and above transaction success.

Overall, the level of variance in success rates decreases for businesses with higher ratings. This demonstrates that a sum-based reputation feedback business rating can be highly accurate for high rating values.

5. RELATED WORK

Research in the field of sociology has examined the ways how embeddedness in social structures have impacted economic transactions. As defined by Mark Granovetter, embeddedness is the theory that economic actions between individuals is so dependent on social relationships that that evaluating behavior based on economic factors alone can be grossly misleading [6]. relationships in economic transactions. Tong analyzed the impact of formal organizational structures on the embeddedness of social relationships [15]. Later work associated the transition to a modern economy with the increase in probability of interactions with partners outside of a narrow relational neighborhood, and attributed it to the increasing small-world effect [9].

A significant amount of literature has looked into the structural properties of social networks including size, density, degree distribution, average distance, and clustering coefficient of these networks. Milgram originally demonstrated the "small world" hypothesis [11], and recent work has proven that small-world networks adhere to power-law scaling characteristics [2].

Online social networks have recently emerged as important areas of study as their popularity has exploded. Kumar et al. study the evolution of structure of the Flickr and Yahoo!360 social networks, and discover that while isolated individuals and communities exist in these networks, there also exists a large strongly connected component spanning the entire network [10]. Others have profiled the power-law, small-world, and scale-free properties of social networks such as Orkut, YouTube [12], CyWorld, and MySpace [1].

Information from social networks can be exploited to enhance many practical areas of research. Mislove et. al. [13] used social networking data to improve Internet search engines, while others have applied this information to increase yields from viral marketing campaigns [4, 14]. Social networks have been proposed as a means to defend reputation systems against Sybil attacks [16], and improve the reliability of reputation systems [7, 8].

6. DISCUSSIONS AND CONCLUSION

While this paper presents a preliminary evaluation of social connections on business transactions, limitations in the dataset prevent a complete analysis of the impact of social networks. For example, the lack of full logs on user interactions means that all social links must be treated as equally relevant, possibly inflating the social connectivity of users using friends as a measure of status [5]. Additionally, privacy concerns prevented us from accessing data on details of user transactions or specific user properties such as length of membership.

As continuing work, we would like to analyze structural similarities between Overstock networks and those typically observed in social networks like Orkut and Facebook. Could any observed differences be attributed to the business nature of the network transactions? We also want to further analyze user behavior on Overstock networks including user lifetimes in these networks, the growth of these networks over the course of transactions, and the relation of transaction value with the personal and business path lengths. Such findings will help us better understand the motivation for users to establish social networks in an e-Commerce setting.

Overstock Auctions is a novel online marketplace that augments traditional auctions with a social networking component. Our study examines the impact of social networking connectivity on business transactions. We find that in general, users perform few transactions with their friends on the social network. However, our sampling of user feedback shows that transactions between partners connected on the social network result in significantly higher user satisfaction. We also find that Overstock's business ratings are accurate, in that they have a strong correlation with user satisfaction after their transaction.

Acknowledgments

We would like to thank Overstock Auctions director Meghan Tuohig for making this paper possible. We thank the anonymous reviewers for their feedback, and Krishna Puttaswamy and Alessandra Sala for their insights and comments. This work is supported in part by NSF CAREER Award #CNS-0546216.

7. REFERENCES

[1] AHN, Y.-Y., ET AL. Analysis of topological characteristics of huge online social networking services. In *Proc. of World Wide Web* (2007).

[2] AMARAL, L. A. N., SCALA, A., BARTHELEMY, M., AND STANLEY, H. E. Classes of small-world networks. In *Proc. of National Academy of Sciences* (2000), pp. 11149–11152.

[3] BHATTACHARJEE, R., AND GOEL, A. Avoiding ballot stuffing in ebay-like reputation systems. In *Proc. of P2PEcon* (2005).

[4] DOMINGOS, P. Mining social networks for viral marketing. *IEEE Intelligent Systems 20*, 1 (2005), 80–82.

[5] DONATH, J., AND BOYD, D. Public displays of connection. *BT Technology Journal 22*, 4 (2004).

[6] GRANOVETTER, M. Economic action and social structure: the problem of embeddedness. *American Journal of Sociology 91*, 3 (November 1985), 481–510.

[7] HOGG, T., AND ADAMIC, L. Enhancing reputation mechanisms via online social networks. In *Proc. of ACM Conference on Electronic commerce* (2004).

[8] JENSEN, C., DAVIS, J., AND FARNHAM, S. Finding others online: reputation systems for social online spaces. In *Proc. of CHI* (2002).

[9] KALI, R. Social embeddedness and economic governance: A small world approach. Tech. rep., University of Arkansas, November 2003. `http://ssrn.com/abstract=462560`.

[10] KUMAR, R., NOVAK, J., AND TOMKINS, A. Structure and evolution of online social networks. In *Proc. of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (2006).

[11] MILGRAM, S. The small world problem. *Psychology Today 6* (1967), 62–67.

[12] MISLOVE, A., ET AL. Measurement and analysis of online social networks. In *Proc. of IMC* (2007).

[13] MISLOVE, A., GUMMADI, K., AND DRUSCHEL, P. Exploiting social networks for internet search. In *Proc. of HotNets* (2006).

[14] RICHARDSON, M., AND DOMINGOS, P. Mining knowledge-sharing sites for viral marketing. In *Proc. of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (2002).

[15] TONG, Q. Embeddedness of social networks in formal organizational structures. In *Meeting of the American Sociological Association* (Atlanta, GA, August 2003).

[16] YU, H., KAMINSKY, M., GIBBONS, P. B., AND FLAXMAN, A. Sybilguard: Defending against sybil attacks via social networks. In *Proc. of SIGCOMM* (Pisa, Italy, September 2006).