MIT Libraries | DSpace@MIT

MIT Open Access Articles

Lending behavior and community structure in an online peer-to-peer economic network

Massachusetts Institute of Technology

Lending Behavior and Community Structure in an Online Peer-to-Peer Economic Network

Katherine Ann Krumme
Media Laboratory
Massachusetts Institute of Technology
Cambridge, MA
kak@mit.edu

Sergio Herrero
Civil and Environmental Engineering
Massachusetts Institute of Technology
Cambridge, MA
sherrero@mit.edu

Abstract— **Increasingly, economic transactions are taking place over social networks. We study the static and dynamic characteristics of a peer-to-peer lending network through 350,000 loan listings and accompanying member profiles from the online marketplace Prosper.com. Our results imply that social factors such as participation in affinity groups and descriptive profile text are correlated with financial indicators; at the same time, we see evidence of suboptimal lending decisions, minimal learning, and herding behavior in the network. We discuss implications and suggest possible improvements to the online peer-to-peer lending model.**

Keywords-economics; social networks

I. INTRODUCTION

Economic transactions have long taken place over social networks; in recent years, an increasing number of models have facilitated borrowing within groups rather than from centralized lenders. Microfinance has proven viable where small-group ties enforce reputation and default rates are minimized [1]. Is has not yet been determined whether reputation and social ties sufficiently strong in an online network of strangers to make peer-to-peer lending a viable model, both from the perspective of borrower access (risk assessment by lenders) and that of loan repayment (expected utility of lenders).

In this paper we characterize the network of a peer-to-peer lending marketplace, in which edges describe a specific economic relationship between two agents, rather than a connection that can be created without (or with low) cost, risk, or reward (as is the case for link-creation in many online social networks). We consider 350,000 loan listings, with accompanying bid histories and member profiles from the online lending network Prosper.com [2]. In addition, we study the co-occurring social network (of online "friendship" and affinity groups), where links are costless but signal a measure of reputation over the economic network. We present simulations of different scenarios of herding behavior and reciprocity in attempt to explain the observed bidding patterns of lenders. Understanding the static and dynamic characteristics of a peer-to-peer network can yield insights about human behavior, as well as serve to improve the model for lenders and borrowers.

II. BACKGROUND

A. Peer-to-peer Lending and Economic Networks

Prosper.com [2] is an example of an online platform for peer-to-peer lending; others include Zopa, Comunitae, and Kiva [6,7,8]. As part of the Prosper marketplace, borrowers create personal profiles and solicit loans via online listings detailing the amount requested, maximum acceptable interest rate, and purpose for the loan. In turn, lenders assess and can bid on listings; if the total dollar amount of bids is equal to the amount requested, the loan is materialized and the funding credited.

In cases when total bid amount exceeds the amount requested, those lenders electing the lowest interest rates are granted a stake in the loan. If a listing fails to garner complete funding, it is canceled by the system and the borrower has the option to repost. Peer-to-peer lending sites collect a percentage of every fully funded loan as it is repaid to the lenders.

The borrower's profile includes independently verified information on his credit history, income, and current debts; a borrower can also elect to join interest-based groups (tied by geography, common interest, or common loan purpose), or collect the endorsements of friends and group leaders.

B. Social Network Analysis

Although scientists have considered network structure since Euler's Konisberg bridges, and social network structure since at least Milgram[9], the analysis of online social networks is a relatively new field.

A recent paper considers qualitative and quantitative methods for segmenting communities of college students participating in the Facebook.com social network [11]. Researchers employ pair counting methods to compare several instances of network partitions, and find that in some but not all cases the quantitative approach results in the same segmentation as the qualitative method of examining other community features (such as major, housing unit, or gender).

In the field of organizational behavior, Krackhardt studies a network of advice-seeking among managers in a small manufacturing firm and finds that influence weights are highest when those advised have few advisors [12]. Similarly, Aral and colleagues find that productivity of executive recruiters can be

predicted from the network structure of within firm as well as external communications [13].

Ryan et al. [4], using a single-variate regression analysis, find that membership in a prosper.com group (part of the social network) increases a borrower's chances of being funded by a factor of 3, relative to the effect of a single-step increase in credit score. From this perspective, the peer-to-peer social network is integral to the network of economic transactions.

C. Experimental Work on Human Behavior Over Networks

A number of economists and psychologists have identified the effects of herd behavior in marketplaces; notably, Robert Schiller describes its effects in the housing market [5].

Salganik et al [10] study the effect of the opinions of others over a large music-sharing network. The authors find a significant effect of dispersion when the number of times a song has been downloaded is listed along with objective information about its artist and genre, empirically highlighting the dynamics of social contagion.

III. METHODS

A. Prosper.com Data

Prosper.com's data is available through the site's API in the form of a multi-gigabyte XML file. The data was imported to a MySQL relational database in order to increase ease-of-use for querying and aggregating information. Prosper's data is divided into nine tables; each member, loan, and group has a unique key which can be used to cross-reference tables. Table 1 offers an overview of the dataset used (consisting of 36 months of data through November 2008).

B. Network Analysis

The network workbench tool [3] was used to render visualizations of subsets of the Prosper economic (bidding) and social (friendship) network, as well as of the networks of individual groups. The circular and spring algorithms based on the Java Universal Network/Graph Framework were used for display.

For each of the networks above, we calculate a clustering coefficent, using the Watts-Strogatz equation as well as the degree distribution for each of the networks.

C. Simulations

Algorithms for the simulations of reciprocity and herding in lending behavior were constructed in MATLAB. We test the outcome given different levels of herding (non-independence of bids) by allowing for user-adjusted biasing of bidding to favor already-supported loans. All independent bids were chosen by constructing matrices using MATLAB's uniformly distributed pseudo-random number generator using the function 'rand'.

IV. RESULTS

A. Network Characteristics

While a small number of borrowers in the peer-to-peer marketplace go on to support the network by acting as lenders, we find an extremely low clustering coefficient in the network of economic links. A visualization of a subgraph (figure 3 in Appendix 1, using Network Workbench [3]) finds that most

TABLE I.   OVERVIEW OF PROSPER.COM LOANS AND MEMEBERS

* 342,341 Listings (requests for loans), including:
o 29,013 which became loans
o 191,251 which expired unfulfilled
o 115,116 of which were withdrawn
* 898,444 members:
o 185,153 registered as borrowers,
o 73,637 registered as lenders,
o 26,804 registered as both.
* 4,365 loans "paid" as of 11/08
* 3,852 loans defaulted, charge-off, or repurchased as of 11/08

components exhibit a tree-like structure. However, the network of friendships and shared affinities shows power-like degree distributions and greater clustering, although not on the order of that observed in other online social networks.

To find instances of greater clustering, we test the graph structure and calculated the clustering coefficients, degree distribution, and average distance of several of Prosper.com's affinity groups. Each group was treated as a subgraph of the prosper friendship graph, with nodes representing members of the group and edges the existence of a friendship link between two members.

Surprisingly, individual groups exhibited very low clustering and few links relative to the number of nodes (although average connectivity was higher than over the entire dataset). In instances of higher connectivity, high link volume was often explained by one highly connected member (typically the group leader) and a star-structured graph. Fig. 1 shows the architecture of a subgroup; Fig. 2 is a visualization of the entire network
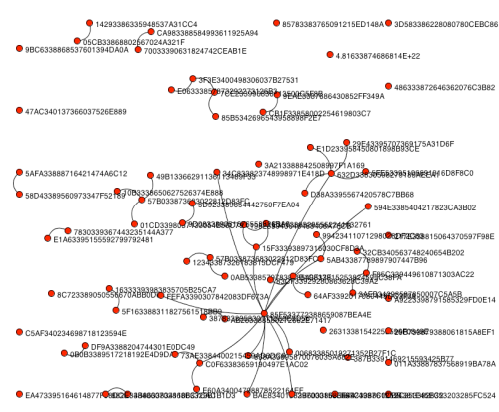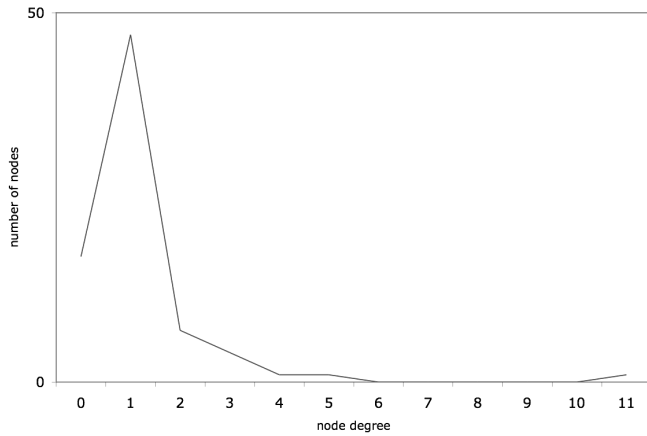
Figure 2. Visualization of Prosper.com social network

Figure 1. Structure and degree distribution of Prsoper.com group 'midwestfriends'

B. Human Economic Behavior

1) Herd Behavior

Social influence can manifest itself between members who have no direct link, either through a lending partnership or by shared group membership. A lender's decision to bid on a listing, for example, may be based on the implicit support of the borrower by other marketplace members. To test this hypothesis, we examine the rate at which bids are accrued to listings (The Prosper marketplace is such that lenders bid incrementally on a listing; if a listing fails to garner complete funding, it is canceled by the system).

From the bid history of completed loans, we find that percentage funding asymptotically approaches the maximum, but that initial funding occurs slowly and then accelerates once an initial number of "pioneer" bids are placed. A large number of bids (9.75% of total) also occur on a small percentage of fully-funded loans as lenders bid down the interest rates of the best loans. Fig 5. shows, for a sample of loans, the percentage occurring in the first, second, third, and fourth sub-periods of the total bidding period (divided equally by time). The effect of herd behavior (as well as of bidding behavior over the graph, such as reciprocity) is further explored in the 'Simulations' section below.

1) Sub-optimal lending (within the network)

One of the aims of online peer-to-peer lending is to allow the "underbanked" and those with low credit ratings to build a reputation that can extend from this marketplace to others. However, analysis finds that lenders maximize their expected by payoff by lending only to those borrowers in the highest credit brackets: that is, credit score is one of the best single features (though not in conjunction with others) for predicting loan repayment (as determined by greedy and forward-backward feature selection methods), and interest rates in the lowest credit brackets are not sufficiently high to make the risk-opportunity tradeoff maximizing of revenues (of course, it is possible that lenders are balancing their risk with investment outside of Prosper.com, or that their utility function comprises of a weight for the "social good" of lending).

Fig. 6 shows the distribution of interest rates versus the percentage of loans that are paid back in full, by borrower credit grade. The expected payoff of investment in a lender with a credit score of AA is higher than that of an investment in a lender with score E, as nearly half of borrowers with score E default (each loan payment is proportionally distributed to lenders as the borrower pays; thus, all lenders will receive some amount of repayment up to the point that default occurs). However, only 18.06% percent of total bids are made on loans with AA borrowers, signaling a failure to maximize expected revenues by lenders in the aggregate. In addition, a sample of lenders shows the distribution of borrowers supported is characterized by high variance (Fig. 5) The only element that serves the underbanked, currently, is the imperfect diffusion of information to lenders.

2) Borrowers Don't Learn

A large portion (91%) of listings never garner sufficient bids to convert to loans. We examine the behavior of would-be borrowers who repost a listing after it expires unfulfilled, and find that (of the minority that reposts rather than exiting the network), 70% re-list only once, and that fewer than 2% of borrowers lower the amount of their request.
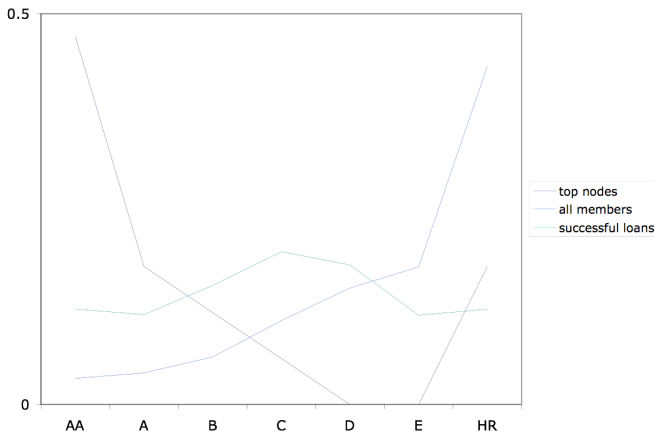
Figure 3. A random sample of best-connected nodes is characterized by a distribution of stronger credit scores

In a marketplace with near-perfect information, we would expect to see greater evidence of learning. Because of the large variance in lending behavior, changes in "subjective" features (the non-financial factors under the most immediate control of borrowers) would, on average, lead to eventual success in loan funding.

Although further textual analysis of listing descriptions and surveys of borrowers are need to lend additional insight to this finding, we see the above results as preliminary evidence that borrowers are either not learning (that is, not adjusting their interest rates and requested amounts to levels appropriate given their financial history), or that there is something endemic in the marketplace (a poor search function, for example) that prevents some borrowers from being matched to the lenders who would prefer their given profile.
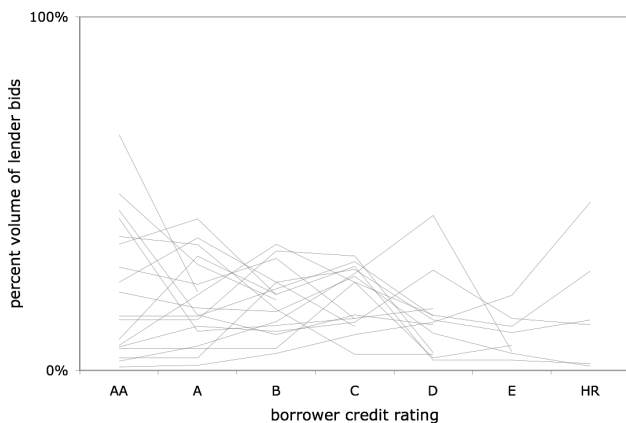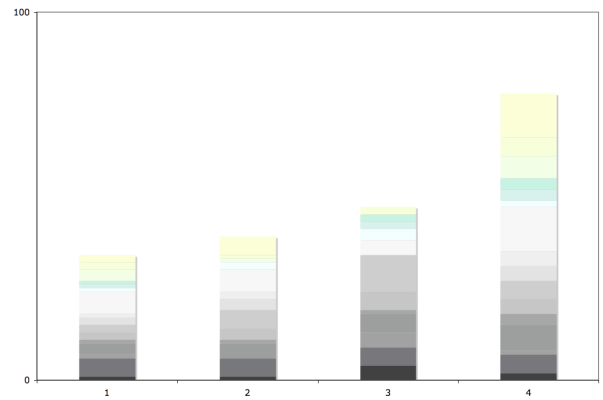

Figure 5. For individual loans and on the aggregate, bidding tends to accelerate in the third and forth quarters. Each bar represents the total bids in each of four equal-time quartiles, for a sample of completed loans

3) Social Features Matter to Economic Performance

Participation in the social network, especially as a member of a group with a high reputation (as determined by Prosper on a scale of 1 to 5), increases the probability of both loan conversion (by almost three times, 17.7% probability versus 6.1% for no group) and of loan repayment (by 1.9% percent late versus 31.1% late for no group). Other analysis [4] has found similar effects for participation in the social network.

In addition, nodes that are highly connected (d > 20) in the social network tend to send more positive economic signals and to fare much better in the economic network. Fig. 3 shows the distribution of credit scores of highly connected borrowers, compared with those of all borrowers as well as the subset of borrowers that ultimately receive funding for a loan. The patterns for other economic indicators (such as debt-to-income ratio, total delinquencies, homeownership) tend to mirror those of credit score.

More work is needed to understand exactly how the social and economic networks affect one another. Is participation in the social network merely a signal of economic health to lenders, or does active interaction with Prosper groups and friends actually help a borrower to meet his commitments?

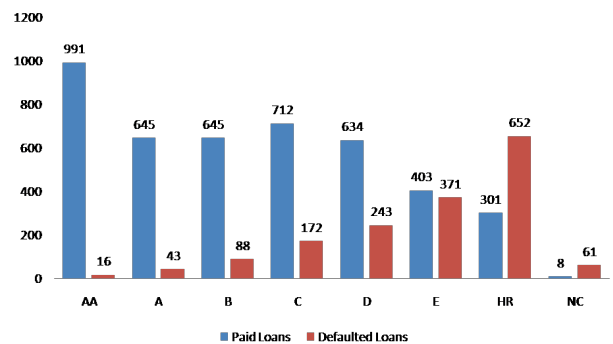

Figure 4. Variance in sample of lenders' beahvior

Figure 6. Expected ratio of paid to defaulted loans by borrower credit grade: even with increasing interest rates for high-risk borrowers, lending to below-B borrowers is sub-optimal

C. Simulations

To better understand the dynamics of bidding interaction, we create a series of simulations based on an interaction model with network structure. The algorithm, coded in MATLAB, allowed the flexibility to adjust the number of bidders involved and the number of bids needed to reach completion, although for the sake of simplicity in the following examples these thresholds were set at 20 and 30, respectively. Each bid was counted as a unit (in contrast to the prosper.com platform, in which bids can take on different dollar amounts. In principle, however, the vast majority of bids are in the $100 range). In addition, the algorithm accepts the following variables

- Number of lenders bidding in each round

- Two or more lenders can bid on the same borrower (Y/N)

- Distribution of herding (right or left or no skew, mean multiplier for already-ahead borrowers)

- Level of reciprocity (in the with-reciprocity model)

Bidding rounds are treated as discrete steps. In each period, two matrices were updates: a) the preference matrix for selected recipients of herding behavior and reciprocity and b) the "results" matrix of total accrued bids. Fig. 7 shows the results of several simulations; in short, we observe that even small multipliers (order of 1.01) for initially-favored borrowers can quickly lead to runaway bidding that might explain, in part, the success of only a small percentage of borrowers.

The effect of the herding begins to become noticeable (that is, loans to which bidders herd accrue, by the end of all periods, a total of at least three standard deviations of the number of bids greater than the mean of the 20 loans at the end of the game) at the threshold at which herding coefficient is as low as 0.1. Below 0.1, and with greater than 30% of loans being thus favored, the effect is not noticeable.

Figure 8 shows the results of a sensitivity analysis of the model [20 loans, 5 of them favored, one with 0.2 boost, one with 0.15 boost, and three with 0.1 boost], including the mean (top) and standard deviation (bottom) of bids after each of 20 bidding rounds: the model is more robust after about 10 rounds, but taken to twenty there is three times as much uncertainty not merely in who will win but in by how much he will win.
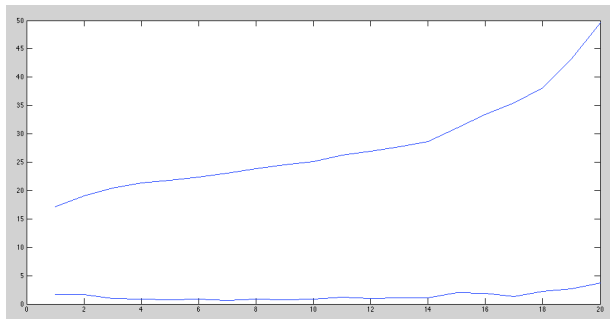

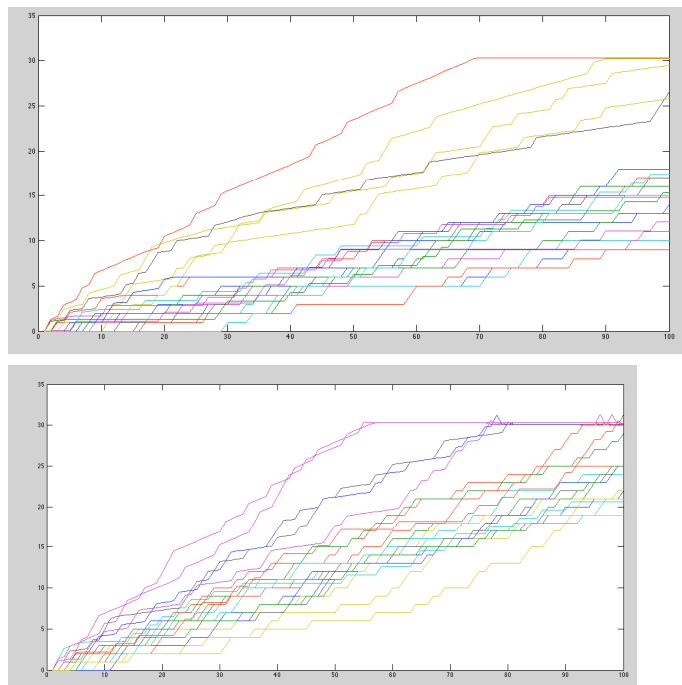Figure 8. Sensitivity analysis of bidding simulation

Figure 7. Simulations of bid aggregation patterns under different levels of herding. Each ribbon represents one loan's pattern of bid accrual; the Y axis is the number of total bids, while the X axis shows each discrete round of bidding. Top image has 3 bids/period and herding coefficients of 0.3, 0.2, 0.1, 0.1, and 0.1. Bottom image has 5 bid/round and coefficients of 0.3,0.3, 0.1, 0.1, and 0.1

V. DISCUSSION

The set of descriptive and empirical results suggest that the peer-to-peer marketplace serves neither borrowers nor lenders optimally; in addition, effects such as herd behavior and poor information dispersion can be observed over the network. At the same time, Prosper's social component serves an interesting (if not fully realized) function as an information signal to prospective lenders and, possibly, a platform for the encouragement of good borrower behavior.

Further work is required to lend resolution to these preliminary results: we do not pretend, for example, to offer any more than a sketch of the motivations of borrowers and lenders in the network. A survey of repeat posters, for example, would help to clarify why borrowers fail to change the amount of a rejected loan. We also hope to amplify our analysis of the network structure of successful groups with a broader statistical measure of the connectivity and performance of all groups participating in the marketplace.

Decision-making in peer-to-peer networks differs from that in lending models at banks; in the former model, lenders have access to subjective factors (e.g. loan description) as well as the judgment of others (through observation of bid dynamics). The

above work finds that lenders make these decisions poorly, that there exists little learning by borrowers, and that the network exhibits low economic clustering despite modest clustering in the social network. These results point to existing shortcomings in the online lending model, as well as to a number of structural changes that could enable more optimal transaction patterns in the future.

REFERENCES

[1] S. Knack, P. Keefer, 1997. Does Social Capital Have An Economic Payoff? A Cross-Country Investigation, Quarterly Journal of Economics.

[2] Prosper Marketplace. http://www.prosper.com/

[3] NWB Team, 2006. Network Workbench Tool. Indiana University, Northeastern University, and University of Michigan, http://nwb.slis.indiana.edu

[4] J.Ryan, K.Reuk, C.Wang, "To Fund Or Not To Fund: Determinants Of Loan Fundability in the Prosper.com Marketplace.", Stanford Graduate School of Business.

[5] R.Schiller Irrational Exuberance (2000)

[6] Zopa. http://uk.zopa.com/

[7] Comunitae. http://www.comunitae.com/

[8] Kiva www.kiva.org

[9] J Travers and S Milgram. 1969. "An Experimental Study of the Small World Problem." Sociometry, Vol. 32, No. 4, pp. 425-443

[10] MJ Salganik, PS Dodds, and DJ Watts, 2006. Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market. Science 311 (5762), 854.

[11] A Traud, E Kelsic, P Mucha, M Porter, Community Structure in Online Collegiate Social Networks

[12] D. Krackhardt, 1987. "Cognitive Social Structures," Social Networks 9:109-134.

[13] Aral, Brynjolfsson & Van Alstyne, 2007. "Productivity Effects of Information Diffusion in Networks" International Conference on Network Science 2007

[14] AP Kirman, 1987. "Ants, Rationality, and Recruitment" Quarterly Journal of Economics 108(1):137-156