Crowd-sourced Financial Support: Kiva lender networks

Gaurav Paruthi, Youyang Hou, Carrie Xu

Table of Content:

Introduction

Microfinance is of critical value as a source of financial services for entrepreneurs and small businesses lacking access to banking and related services. Created in October 2005 as the first peer-to-peer microlending site, www.kiva.org(Kiva) is a platform for low-income entrepreneurs in developing countries to fundraise for their business from other individuals. Through Kiva, anyone can make an interest-free loan of at least $25. As of Nov. 2013, Kiva was able to raise about $1 million every three days.

Here is how Kiva works. After the initial screening, a potential borrower will build an online profile on the Kiva website. A lender will the pick a business and make a loan using their credit card. Then Kiva transfers the loan to the local partners, and partners disburse the loan to local borrowers. Partners will then collect the revenues from borrowers and pay the loans back to the lenders again through Kiva. Previous researchers have investigated how the characteristics of borrowers and lenders influence funding behaviors. For instance, Meer & Rigbi (2013) found that whether to fund a loan or not depends on the language barriers and social distance.

In August 2008, Kiva created “Lending Teams”. Kiva lenders can voluntarily join and organize their own teams. Kiva teams cover a broad range of themes and interests. Team members can communicate through team messages in team forums (Figure 1). Each loan made by the members is counted towards team contribution to Kiva borrowers. Teams are ranked according to the team contribution. The team feature makes Kiva lenders more motivated to participate in lending activities because of the team identity. The team ranking also may foster competitions among different Kiva teams.

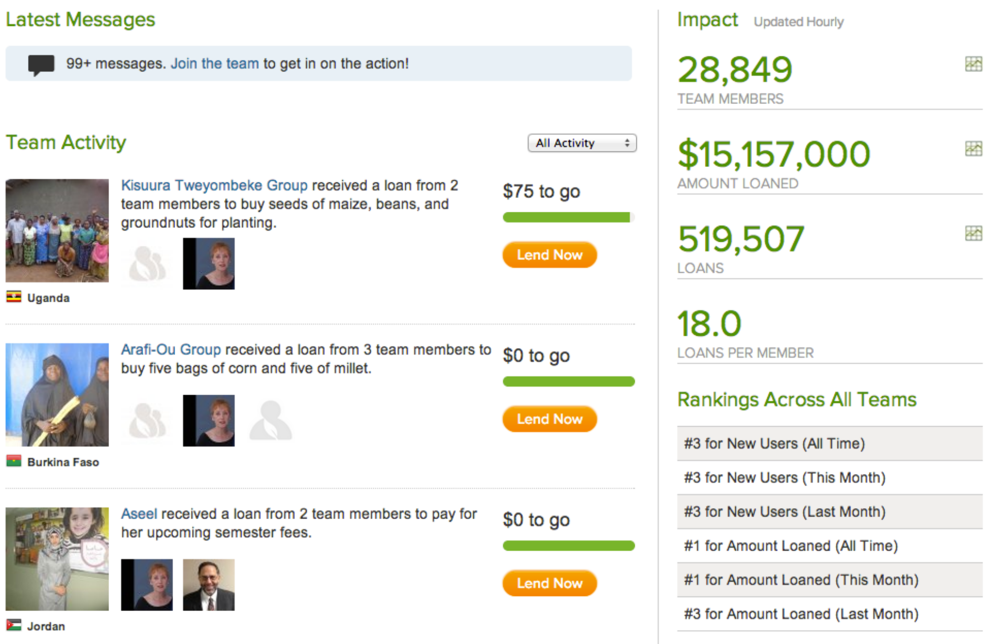


Fig 1. Snapshot of a Kiva team page

Research shows that the team features has an impact on funding behaviors. For instance, Liu et al. (2012) found that one makes more loans after joining a team. Chen et al. (2012) further explored the causal effects of team forum messages on nudging team members' lending decisions. By manipulating the message contents in Kiva team board, they found that forum messages increase lending from team members. Messages with social information about coordination and competition encouraged more team lending.

Drawing on the previous literature about the effects of Kiva teams on lending behaviors and outcomes, this study looks at how the lender network of Kiva based on the team membership correlates with the fulfillment of loans and the outcomes of the fulfilled loans. In particular, we are interested in how lender network structures correlates with (1) how soon a loan is funded, as well as (2) whether the loan gets funded by the deadline.

Method

To answer our research questions of understanding how soon a loan is funded or whether it gets funded or not, we brainstormed and came up with the idea that looking at the lender network for each loan could provide us some good insight. We believed that the use of network concepts could be very beneficial in answering these questions.

We proposed multiple lender networks for each loan, where the edge between two lenders in each network would signify that the lenders share a team, location, gender and join date. Amongst these we finalized team and location networks, as we think that they would affect the duration of a loan funding the most. We do this based on our intuition that sharing a team or location might accelerate the loan funding. We base this on the finding that people in teams tend to lend more (Liu et al., 2012). We also think that, gender and join date would be interesting but decided to analyze them in future work.

For retrieving the actual data, we leveraged two sources provided by Kiva. Firstly, Kiva provide a snapshot[1] of the most recent information on the loans, lenders and which lenders loaned to which loan. Some statistics can be found in Table 1. Next, for lender team information, we used the API[2] provided by Kiva. We gathered team information for about 1,200,000 lenders on kiva.

The implementation of the data gathering scripts was on Python. We used IPython Clusters [3] to significantly improve the speed of data gathering. This process continued for many days as the Kiva has a strict limitation on data rate and throttling. We stored the data and results in a MongoDB server for on-demand retrieval.

Table 1. KIVA Snapshot data stats

Number of Lenders	1234204
Number of Loans	670764
Average loan amount	932.30
Average number of lenders donating to a loan	$27.42
Average time taken to fund loans	2.5 days

For the network creation, we iterated through each loan and retrieved the list of lenders that contributed to that loan. We then created the location and teams networks. The location network was created in a fairly straightforward way by comparing the location of all pairs of lenders. For the teams network, we find the pairwise intersection of the teams that each

[1] Kiva snapshot: http://build.kiva.org/
[2] Kiva API: http://build.kiva.org/api
[3] Using IPython for parallel computing, http://ipython.org/ipython-doc/dev/parallel/

lender is a member of. The edge weight between two lenders denotes the number of teams shared between the two lenders.

For each network, we then calculated the density, average clustering, number of connected components, number of nodes, the number of edges and the number of weighted edges. We decided upon these measures based on our intuition about what these measures represent. The Density of the network represents how important was a particular team or location for a given loan. Average clustering for teams would represent how closely the lenders are to each other irrespective of their team membership. Number of connected components represents the communities in the network and would give us a measure of the diversity.

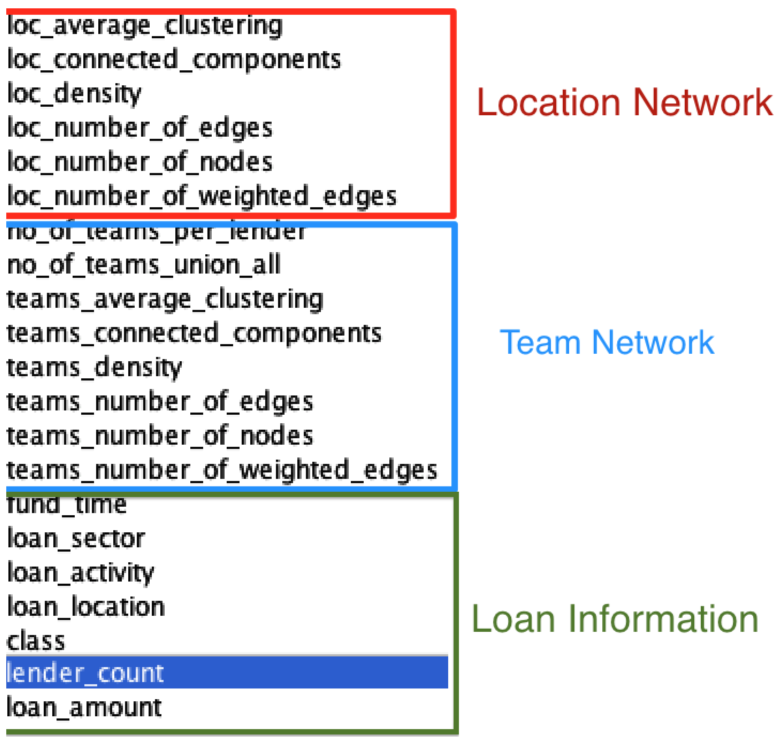


Fig 2. Final data that we created for analysis

Findings

Table 2. Lender network structure characteristics

Network	num of nodes	num of edges	num of components	average clustering coefficient

Team	5.345	9.07	1.778	0.175
Location	24.32	162.1	10.1	0.517

First, we examined how the network structure influence the eventual outcome of a loan: whether it is funded or not funded. We ran a logistic regression to predict the loan outcome using the tie-ratio, team clustering and number of connected component in the lender network. We define the tie-ration for a loan as the ratio of the number of teams shared between lenders and the total number of teams for a given lender. We also controlled the total number of lenders as well as the amount of each loan. The results showed that tie-ratio and team average clustering coefficient are negatively related to the loan outcome. However, the number of connected component of lender team network positively predicts the loan outcome.

```
                           Estimate Std. Error z value Pr(>|z|)
(Intercept)                15.03157    1.32689  11.328  < 2e-16 ***
log(loan_amount)           -4.99038    0.20452 -24.401  < 2e-16 ***
log(lender_count)           7.21794    0.26519  27.218  < 2e-16 ***
tie_ratio                  -8.22142    0.94199  -8.728  < 2e-16 ***
teams_average_clustering   -2.39808    0.65686  -3.651 0.000261 ***
teams_connected_components  0.13544    0.05851   2.315 0.020625 *
```

Then we tested on the network constructed from the location of lenders who supported the loan. Both average clustering coefficient and number of components are positively related to the success of a loan.

```
                          Estimate Std. Error z value Pr(>|z|)
(Intercept)             -17.240040   0.155390 -110.95   <2e-16 ***
log(loan_amount)          1.105073   0.025234   43.79   <2e-16 ***
loc_average_clustering    7.711800   0.099559   77.46   <2e-16 ***
loc_connected_components  0.196027   0.002364   82.92   <2e-16 ***
```

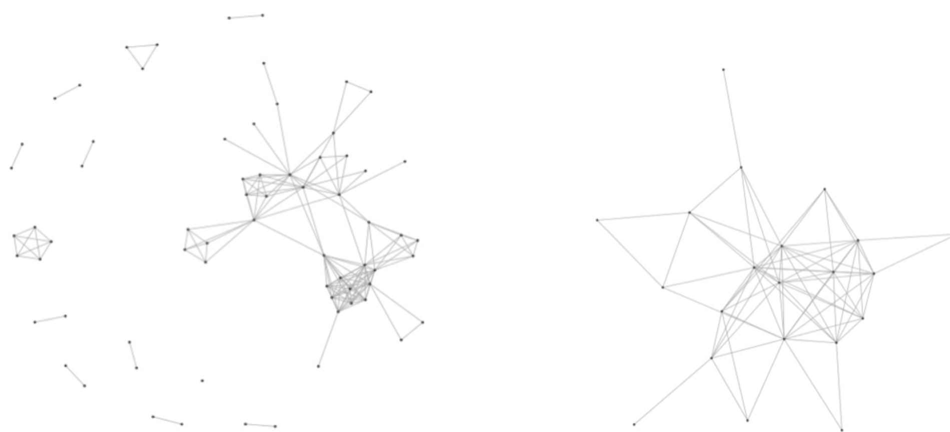
In another set of regression models, we examined how lender network structures influence the speed of a loan got funded. In the first model, we tested how network connected by the team membership influence the speed that a loan gets supported. The tie ratio and average clustering coefficient predict slower speed, whereas the number of connected components predicted faster speed that a loan can get funded.

```
                             Estimate Std. Error t value Pr(>|t|)
(Intercept)                -14.527423   0.287427 -50.543  < 2e-16 ***
log(loan_amount)             1.312436   0.035021  37.476  < 2e-16 ***
tie_ratio                   -3.268018   0.237758 -13.745  < 2e-16 ***
teams_average_clustering    -1.251035   0.141209  -8.859  < 2e-16 ***
teams_connected_components   0.031323   0.009423   3.324  0.00089 ***
```

Finally, we tested the effect of location network structure on the speed of a loan got funded. Similar to the result of loan success, the average clustering coefficient and connected components both predict higher loan funding speed.

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	-15.511135	0.188608	-82.240	< 2e-16	***
log(loan_amount)	1.380433	0.032009	43.127	< 2e-16	***
loc_average_clustering	-0.462850	0.125144	-3.699	0.000217	***
loc_connected_components	-0.014454	0.001525	-9.481	< 2e-16	***

Discussion



High Number of components
Fast
ln_id=488584
speed=0.0036

Low Number of components
Slow
ln_id=524092
speed= 0.00092

Fig 3: Comparison of two networks with different number of components. Higher number of components performs faster than the network with low number of components.

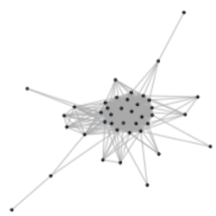
Diversity and Competition

We found that the number of connected components of lender team networks positively predicts the loan outcome as well as how fast the loan is funded. The number of connected components signify the different communities of teams that lend to a given loan. This implies that diversity in the participating lenders helps in reaching the goal of loan. The diversity in the network indicates that lenders who contributed belonged to different teams that do not bridges linking them. Bridge between two teams would imply there is a lender that is a member of both the teams. In the fig X, we show two examples of the network from our analyzed loans.

The answer to why diversity positively affects the loan outcome, we think it could be because of multiple reasons. Competition between teams have shown to motivate team members to lend more (Chen et al. 2012) and could explain this result. Another reason, could be because the competition between loans to have visibility with teams.

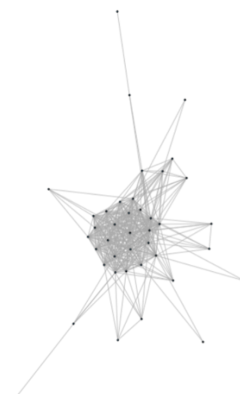
Information diffusion

$$\text{ties\_ratio} = \frac{\text{number of shared teams per lender for this loan}}{\text{total number of teams per lender for this loan}}$$



Weak Tie
Faster

id	lender_count	tie_score	speed
463532	95	0.052632	0.004777



Strong Tie
Slower

id	lender_count	tie_score	speed
551810	67	0.508876	0.000984

Fig 4. Weak ties help in funding the loans faster. Note that the ties do not affect the network structure here. The tie score characterises the whole network. A given lender network could have a weak tie score if its a network of lenders that are weak-ties in the grand lender network of kiva.

We analyzed how ties between lenders in a network affect the outcome of the loan. We defined the measure tie score for a loan as the ratio of the number of teams shared between lenders and the total number of teams for a given lender.

We observe that tie score negatively affects the speed. This implies that if a network characteristic of weak ties then the loan funding speed was higher. We draw the analogy of finding from the paper "The Strength of Weak Ties" (Granovetter, 1973), where it was shown that weak ties are better in getting information about job opportunities. Similarly, we think that information about loans follows a similar pattern. If there are costs associated with examining team messages, a smaller tie score means that a pair of connected lenders are more likely to see each others' activities/messages. Therefore, the information about one lender's lending activities is more likely to reach the other one.

Example:

Situation 1: A and B share 10 common teams. Due to time costs, A only checks 2 team forums once in a while. Then A only has 2/10 probability seeing B's activities through the forum.

Situation 2: A and B share 2 common teams. Due to time costs, A only checks 2 team forums once in a while. Then A only has 2/2, 100% probability, seeing B's activities through the forum. Then information about A is more likely to flow to B in this situation, with a smaller tie score.

Future Work and Conclusion

Community detection

Here we only tested the number of components and clustering coefficient of lender networks. Future work could further analyze the sub-communities of lender network, and how different networks influence the performance of each loan.

Evolution of networks

Another way to validate our hypothesis regarding the relationship between the network structure and the loan outcomes is to track the dynamics of the lenders network and test if the construct network will lead to better or worse results. Additionally, future research could look at how the loan or lender characteristics influence the evolution of networks.

Role of bridges in networks

One interesting observation we had on the lenders network is that for some lenders, they kind of bridge diverse lender teams or communities for a loan. For instance, in Figure X, there are four nodes sort of connecting two lender communities together. Who are those lenders? What are their roles in the lender networks? This opens many potential challenges for future research on the Kiva networks.

Reference

- Chen, R., Chen, Y., Liu, Y., Mei, Q. (2012). Social identity in online microfinance: A field experiment at Kiva. http://www.uta.fi/yky/en/studies/disciplines/northamericanstudies/summer_school/2013_Yan_Chen_Kiva_Experiment_2013_06_60min.pdf
- Liu, Y., Chen, R., Chen, Y., Mei, Q., & Salib, S. (2012). I loan because...: Understanding motivations for pro-social lending. In Proceedings of the fifth ACM international conference on Web search and data mining (pp. 503-512). ACM.

- Meer, J., & Rigbi, O. (2013). The Effects of Transactions Costs and Social Distance: Evidence from a Field Experiment. *The BE Journal of Economic Analysis & Policy*, *13*(1), 271-296.
- Granovetter, Mark. "The strength of weak ties." *American journal of sociology* 78.6 (1973): l.