Analysis of An Investment Social Network

Alex Dusenbery Khanh Nguyen Duc A. Tran
Department of Computer Science
University of Massachusetts - Boston
Email: {adusen, knguyen, duc}@cs.umb.edu

***Abstract*—Much research has been done in the last several years on social network dynamics as online social networks have become ubiquitous. While networks like Facebook and LinkedIn have given the research world access to data on large-scale, general human social interaction in an online environment, there are a growing number of smaller-scale social networks aimed toward certain niche user groups which can provide a richer context for the study of social network dynamics. One such network is Currensee, an emerging social network for foreign currency exchange (Forex) traders. Since all of the trades represented on the Currensee network are the actual trades being made by users, it can serve as a good model for the study of financial activity by individuals in the context of a social network. We present in this paper an analysis of this network in terms of its connectivity, social activity, and trading activity.**

I. Introduction

Complex network research has been an active area of study since the inception of the random graph model by Erdős and Rényi [1]. In this model (ER), any two nodes are joined by an edge with some fixed probability. This means that the degree distribution of a random graph approaches a Poisson distribution, where most nodes will have a degree very close to the average node degree in the network. As the study of complex networks has grown, it has been discovered that many real-world networks diverge from the ER model. The degrees of nodes in some real-world networks have been shown to follow a Power-law, rather than a Poisson, distribution. That is, for large degree k, the fraction of nodes in the network having this degree is $P(k) \propto k^{-\lambda}$, where the parameter λ is the power-law exponent for the degree distribution.

In studying complex networks, researchers have become increasingly interested in online social networks, which are web-based networks formed by users in often diverse physical locations. Sites like Facebook, MySpace, and LinkedIn are an extremely popular and convenient way for people to interact, keep in touch, find new friends, or look for employment. As data from many online social networks has become available, it has provided a means to examine some sociological properties of these networks. The online interactions of individuals provide insight into the dynamics upon which human social interaction is built. The process by which individuals in a social network form cliques or exchange information has important implications in epidemiology and marketing, while the online interactions themselves can provide an understanding of the structure of Internet traffic and how to best shape that structure for its growth.

In this paper, we analyze a data set from Currensee [2], an online social network aimed specifically at foreign currency exchange traders. The primary incentive for a person to join the Currensee network is to share their trading activity with their friends in the network. A person can see, in real-time, the actual Forex trades being placed by any of their friends in the network. The Currensee network therefore is very unique, which provides an opportunity to study social network topology in the context of the investment activity of a group of users. This could provide insight into how the trading activity of a group may influence that of an individual, or what the relationship is between the popularity of an individual in a social network and their success as a trader. In this paper, we attempt to answer the following questions (and those related):

- What basic topological properties does the Currensee network exhibit? Does this network share the same topological properties with other social networks? For example, does it have a power-law degree distribution?
- What are the distribution of user trading activities and that of social activities such as sending and receiving private messages? How actively a user log in the website?
- What is the relationship between a user's social popularity (their degree) and their trading activity? Do socially popular users in the network tend to be the users trading most frequently? Do they tend to be the users who are making the most money on their positions?

In examining the first two points, we'll find properties somewhat similar to other online social networks. In examining the third, however, we will come across something a bit different: the best performing traders in our network are often the least-connected. This raises an interesting question about the role of incentive to share information in social networks.

This paper is organized as follows. The related work is reviewed in Section II. The data set and some preliminaries are discussed in Section III. The results of our analysis are presented in Section IV. The paper is concluded with our directions for future work in Section V.

II. Related Work

An early paper on online social networks is [3] which studies the Club Nexus website of Stanford Universty. Yahoo 360 (now defunct) is analized in [4]. A series of work on popular online social networks, namely Orkut, Flickr, Youtube and Facebook, is reported in [5]–[8]. The general finding is that each of these networks exhibit small-world and scale-free properties, somewhat consistent with the preferential attachment model [9]. The evolution of activity inside a social network has also been investigated. For example, the activity network of Facebook and Cyworld social networks is analyzed in [6] and [10], respectively. In a different effort, the authors of [11] study user communication patterns in an Instant Messenger network, using a dataset of more than 300 billion messages. It is found that a strong homophily exists with the tendency for similar users to interact with each other (based on location, sex, age, etc.).

A good review of exchange rate economics is given by Hopper in [12]. Very generally, the goal of foreign currency exchange trading is to exploit the fluctuations of exchange rates between different currencies in order to make a profit. Hopper's review summarizes some of the early models of exchange rates, as well as the idea that the driving force behind the short-range determination of exchange rates is market sentiment, rather than economic fundamentals (such as money supplies, interest rates and trade balance). This is an interesting light in which to view the study of the Currensee network, as the sentiment (expressed directly by real trading activity) of any individual user in the network can be displayed for the entire network to view. To the best of our knowledge, ours is the first study of an online social network connecting users around their individual investment activity.

III. The Dataset

The Currensee network is an online social network formed by people who trade on Forex markets and who have linked their brokerage account to Currensee. In this way, the actual trading activity of users is displayed on Currensee for the entire network, or often only a user's friends, to view. A user can also participate in many of the traditional activities of online social networks: forming friendship links, sending private messages, and starting or participating in public discussions and polls. She may utilize a public profile to display personal information such as date of birth, location, occupation, phone number, email address, blogs, websites, hobbies, and interests. They can also post different trading strategies that they are using for a given position, track the success of strategies, and view their (or others') trading performance over time.

The data was made available to us directly by Currensee. The data spans the range from the creation of the network in early 2009 to June 1, 2010. The network consists of active 4,802 users and 24,150 friendship links. We consider the time at which a user completed her registration to be the time at which they entered the network. Since our dataset indicates the existence of undirected friendship links, but does not indicate the time at which a link was formed, we consider friendship links to have been formed at the time when one user sent a friend request message to the other user (given, of course, that the request was accepted). In this data set, there are 3,791 private messages sent from one registered user to another. At the time this data set was extracted, there were 1,087,247 positions (trades) linked to Currensee user accounts.

To observe the evolution of the network's properties, we looked at the data in terms of the number of users registered. That is, our time steps were based on number of users registered in the network, rather than dates and times.

IV. Analysis

The purpose of our research is to examine the distribution of user social connectivity, user social activity, user trading activity, and the relationships between these distributions. Specifically, we want to answer the questions raised in Section I. We use the algorithm of [13] for power-law fitting. The results of our analysis are discussed below.

A. General Topological Properties

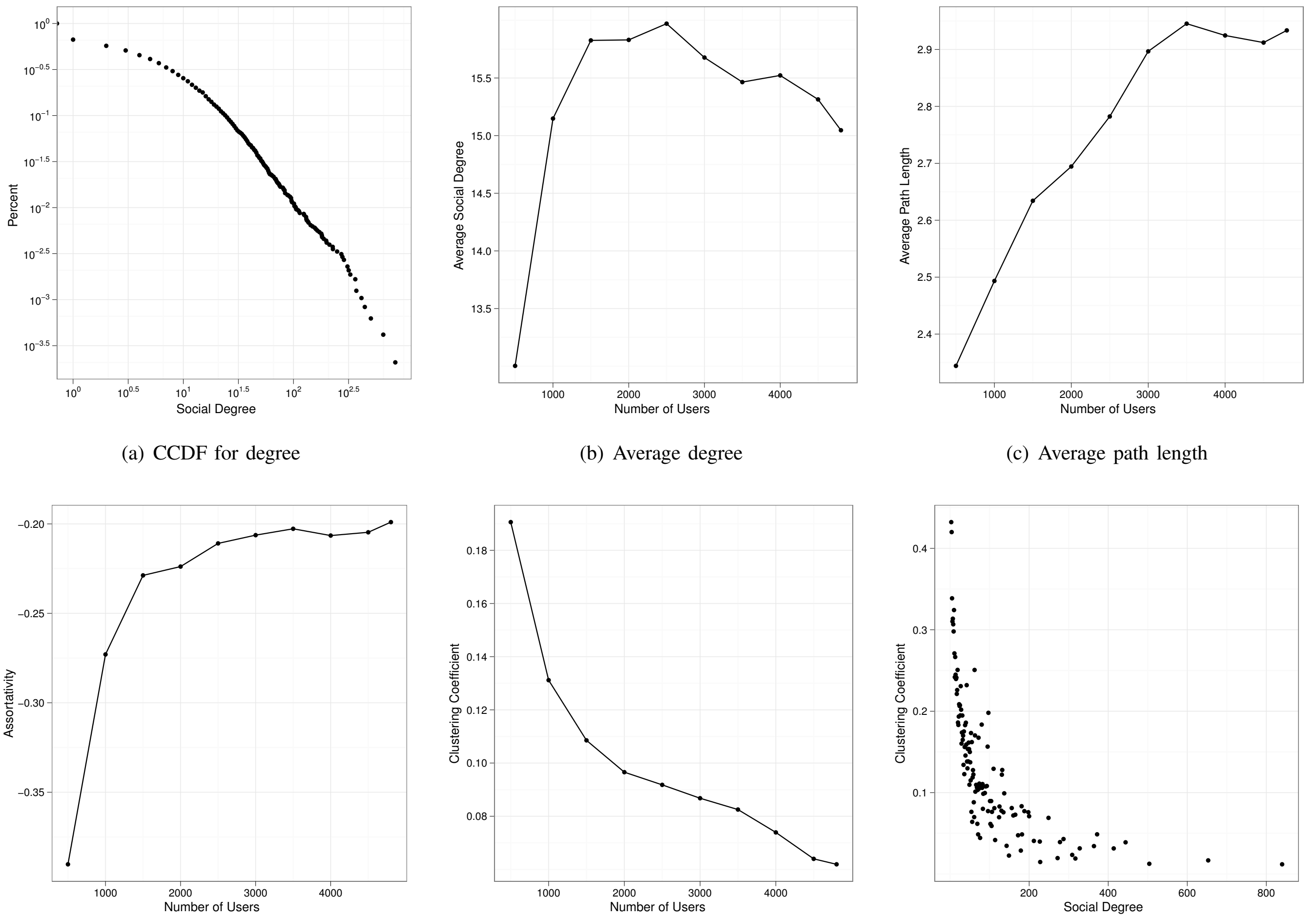
We first examine some of the general topological properties of the Currensee network. As this is still a fairly young network with a relatively small number of users, we expect to see patterns with drastic changes during the network's early formation, giving way to stablization as the network grows.

1) Social Degree: The social degree of a user is the number of friendship links in which she takes part. While the Currensee network is much smaller than many general-purpose online social networks, we still expect to see a power-law distribution for node degree. Figure 1(a) shows the complementary CDF of the social degree distribution. Most users in the network have a small number of friendship links, while a few users have a great number of friendship links. This distribution can be approximated with a power-law distribution with exponent 2.6 (for degrees above xmin = 34, fitting error D = 0.022).

Figure 1(b) shows that the average social degree of a user in the network makes a rather large jump as the network increases in size from 500 users to 1000 users, and then experiences gradual changes. This makes sense when thinking about a social network in its early stages. It suggests that the early users of the network are forming friendship links with other users at an accelerated rate, while as more users enter the network, the early users are less inclined to form friendships with new users. It is too early to tell whether the average degree will converge or keep decreasing as the network keeps growing. If the average does converge, it will be consitent with that observed in some other social networks, such as Facebook [6].

2) Average Path Length: Figure 1(c) shows that the average path length of the network increases steadily until about 3000 users have joined the network, at which point the average path length appears to be stable with the addition of new users into the network. The average path length appears to converge (to approximately 2.9), which is evidence that the Currensee network is very likely a small-world network.

3) Assortativity: The assortativity (or homophily) coefficient of a social network refers to a node's preference to attach to other nodes of similar social degree [14]. As a consequence of this preference, nodes that are highly connected will often form links with other highly connected nodes. Figure 1(d) displays the assortativity coefficient of the Currensee network, which shows a sharp jump between the network's inception and the first 1500 users entering the network, and a more subtle



(a) CCDF for degree

(b) Average degree

(c) Average path length

(d) Assortativity

(e) Global clustering coefficient

(f) Local clustering coefficient as function of degree

Fig. 1. General topological properties

fluctuation thereafter. However, it is important to note that, thus far, the Currensee network shows disassortative mixing, as indicated by the negative assortativity coefficient. This is more reminiscent of technological and biological networks, which exhibit disassortative mixing, unlike the assortative mixing seen in most traditional social networks [15]. An important consequence is that this social network would be easily disconnected by the removal of just a few highly connected nodes from the network.

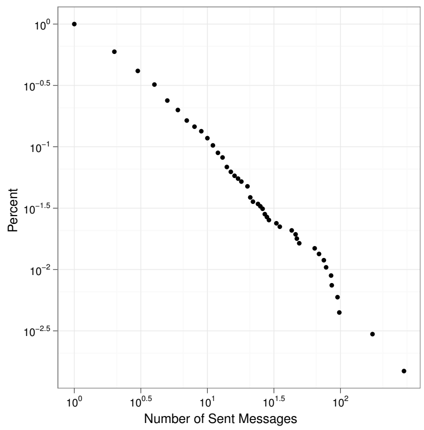
4) Clustering Coefficient: Figure 1(e) displays the global clustering coefficient (or transitivity) of the network. The global clustering coefficient is the ratio of triangles in the graph, whereby two neighbors of a node are themselves joined by an edge. The clustering coefficient for this network experiences a sharp decline in its early stages, followed by a gradual decline as more users join. As the network grows in size, the clustering coefficient tends to reach 0. The shape of such a network becomes a sparse one in which very few triangles are present, because the neighbors of a given node are most likely not connected themselves. In contrast, a network with high transitivity is quite dense, as most neighbors of any given node are themselves connected.

Figure 1(f) shows how the local clustering coefficient of a user varies with its social degree on Currensee. The local clustering coefficient of a user here is defined as the ratio of the number of links that exist between a node's immediate neighborhood and the maximum possible number of links. In this network, users with lower social degrees have higher clustering coefficients. This implies that the network is formed from a number of highly connected hubs, at the boundary of which nodes tend to form tightly connected cliques with their close neighbors.

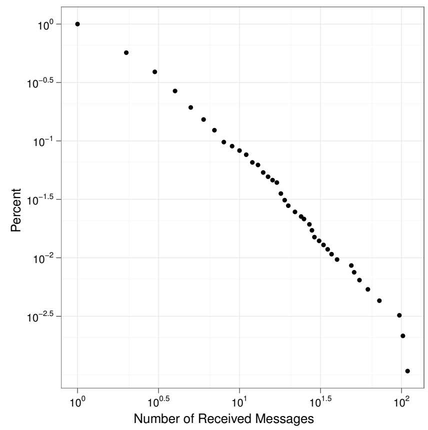
B. Social Activity

We evaluate social activity in terms of (1) how actively a user participates in exchanging private messages and logging in the network, and (2) how long a user stays online in the network. In exchanging messages, a user is not required to have a friendship link to another user in order to send her a private message.

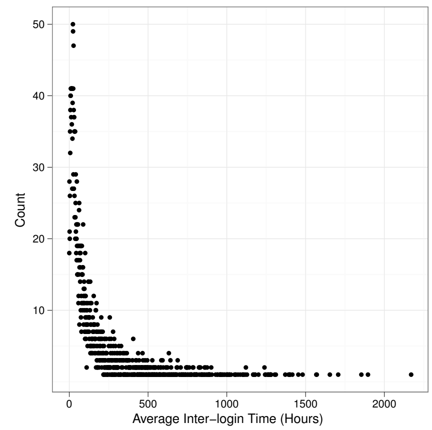
Figures 2(a) and 2(b) show that the number of messages sent and that received by a user follow a power-law distribution,



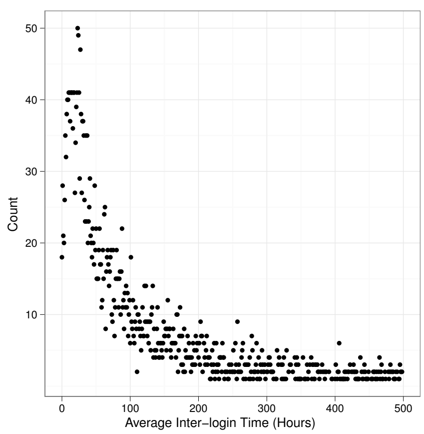
(a) Sent message distribution (CCDF)



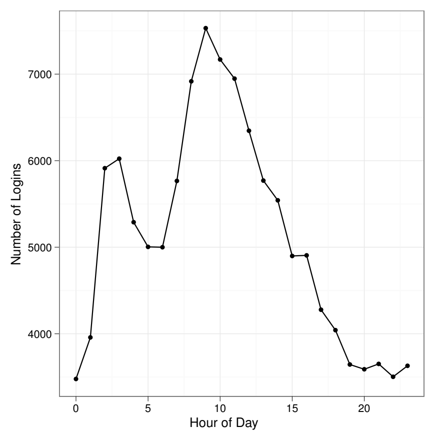
(b) Received message distribution (CCDF)



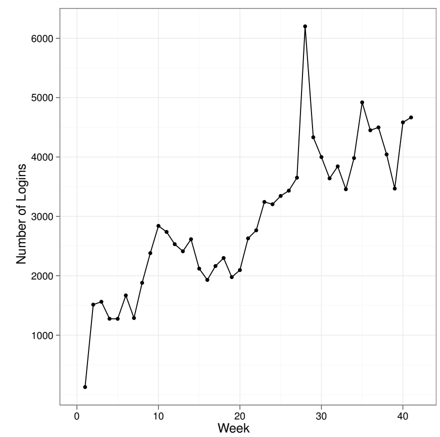
(c) Distribution of avg. inter-login time



(d) Zoom-in of distribution of avg. inter-login time



(e) Number of user logins by hour



(f) Number of user logins by week.

Fig. 2. Analysis of social activity. All times given are Eastern Standard Time.

with power-law exponents 2.31 (xmin = 9, D = 0.039) and 2.24 (xmin =3, D= 0.024), respectively. In comparison, a power-law pattern is also observed in the message activity distribution of Facebook [16].

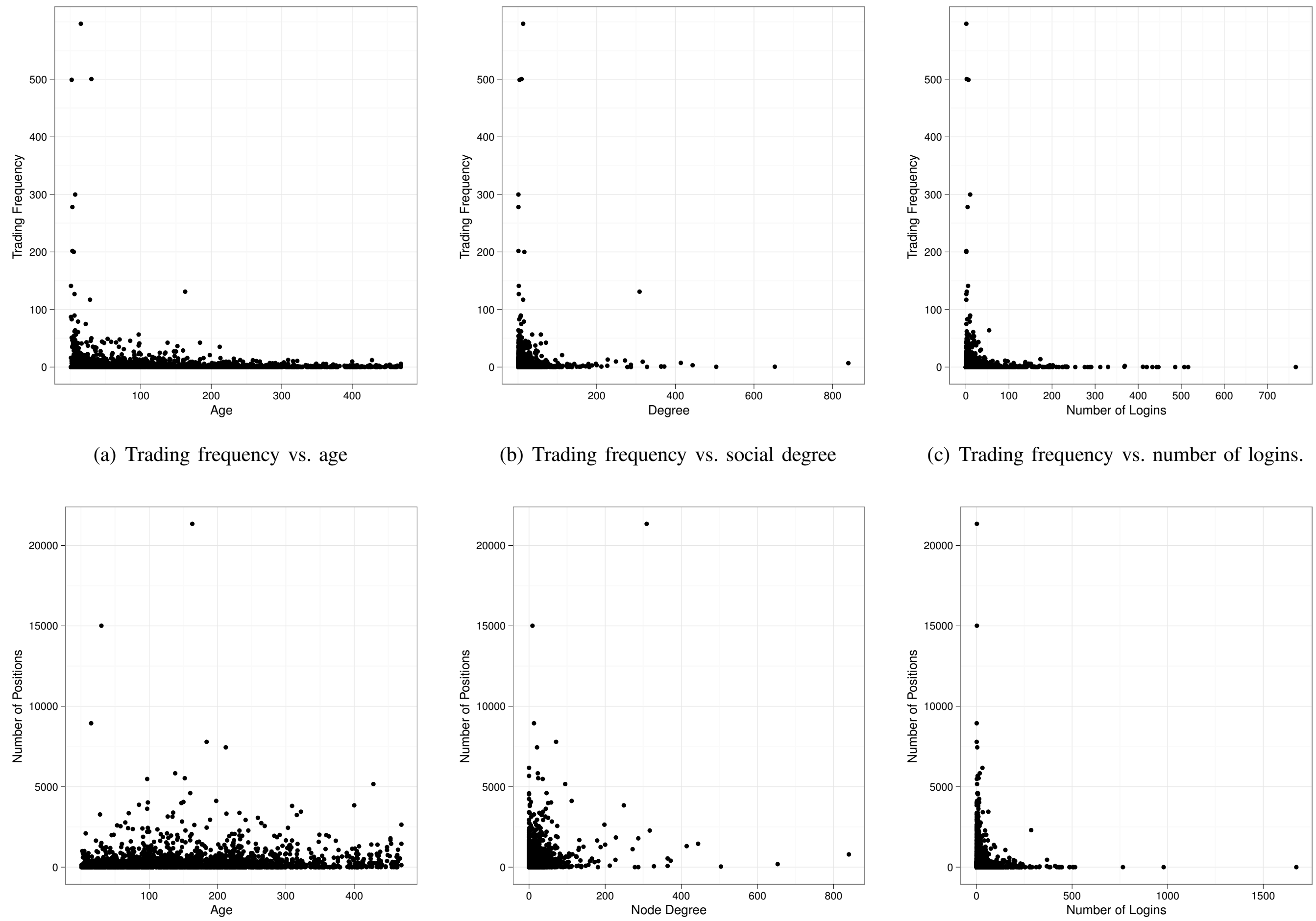
To determine how actively a user logs in the network, we compute the average inter-login time for each user and plot its distribution in Figure 2(c) (and a zoom-in in Figure 2(d)). An interesting observation is that the distribution appears to be log-normal, in which extremly short and long inter-login times are rare. Figure 2(e) illustrates that the site is busiest around 9 A.M. EST, which coincides with the opening of Forex markets in the United States. The spike around 3 and 4 A.M coincides with the opening of markets in the United Kingdom. This simply suggests that users are most active in the network while trading is open, so that they might follow the trading activity of other users in the network. Figure 2(f) shows the correlation between the growth of the network and the general user activity in the network. There appears to be a generally linear increase in site logins as the network grows in size.

C. *Trading Activity*

Currensee is unique in that it is the first social network for Forex traders, and it displays information on the actual trades being made by its users. We are interested in trading activity in this network in terms of trading frequency, number of positions opened, and trading success, at the level of individual users.

We define the trading frequency for a user as the total number of trades they have made divided by their network age (i.e., the number of days as a member of the network). Figure 3(a) shows that long-time members of the network are not neccessarily the most active in the number of trades made per day. Figures 3(b) and 3(c) further imply that nodes that are the least active socially are the nodes with the highest trading frequency. It seems that the most active traders (in terms of the number of trades per day) can be found among those nodes that have low degree and have rarely logged in the network.

The comparison between age in the network and total number of positions is expressed in Fig 3(d), showing no correlation. The likelihood that a user makes a certain number of trades does not depend on age and as such the trading



(a) Trading frequency vs. age

(b) Trading frequency vs. social degree

(c) Trading frequency vs. number of logins.

(d) Number of positions vs. age

(e) Number of positions vs. social degree

(f) Number of positions vs. number of log-ins

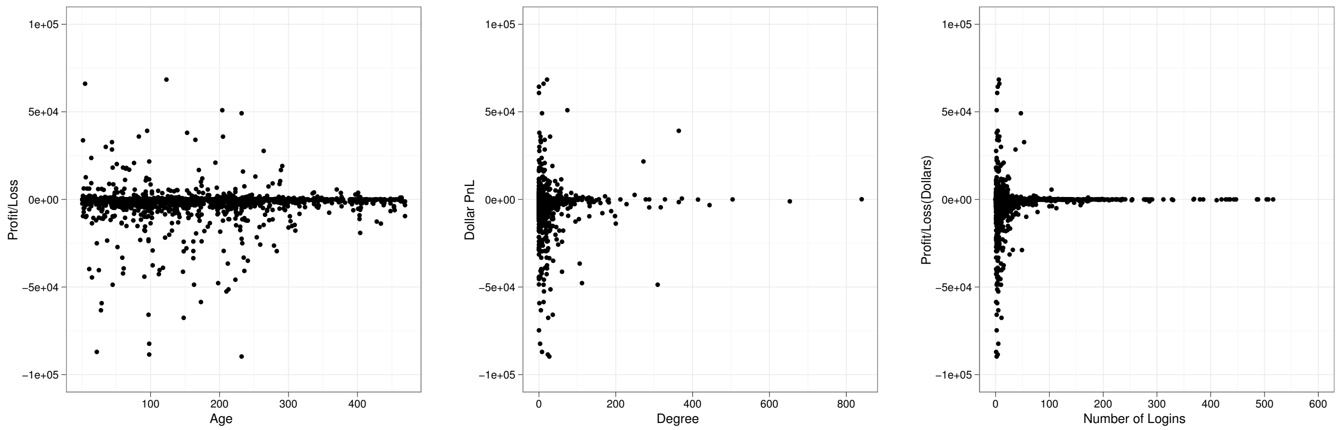
Fig. 3. Trading activity in terms of trading frequency and number of positions opened

frequency of old nodes (long age) should be less than that of young nodes (small age). A young node should also have a few friendship links and a small number of log-ins. Consequently, this could serve as a likely explanation for the trading frequency observations discussed earlier.

As Figure 3(e) indicates, the users who have made the most trades tend to be users with low degree. There are two likely explanations for this sort of behavior: One, those users with high degree but few positions made may be new, inexperienced users who have just joined the network and simply tried to befriend as many other users as possible. These users may be trying to gain experience simply by interacting with and observing the trading strategies of as many traders as they can. Two, the users who have made many trades are probably skeptical of forming friendship links with inexperienced traders. There is no real incentive for an experienced trader to interact with inexperienced traders, with the possible exception of serving as a mentor to one or two new traders. As such, an experienced trader may have a small social degree.

Figure 3(f) demonstrates the relationship between the number of positions opened and the number of log-ins. It is observed that users who made many trades tend to login the least. This raises the question of how interested these experienced users are in participating in the social network. Perhaps there should be some incentive scheme to get these users more involved in the network to make the network more of a value to the other members.

To measure the relationship between the trading success (profitability) of a user and its social connectivity, we looked at the cumulative profit or loss of positions for each user (in dollars) in the correlation with its age, social degree, and number of log-ins. Figure 4 displays the results. As active traders seem to have lower degree, it seems that the traders who are most successful also tend to have lower degree (Figure 4(b)). This may again stem from a lack of incentive: successful traders may tend to be somewhat reclusive in the context of an online social network. Figure 4(c) shows that the users who make the most money are not logging in to the site very frequently (however, it also indicates that the users losing the most money do not login that frequently, either). The users that login most frequently tend to break even in terms of trading



(a) Profit vs. age

(b) Profit vs. social degree

(c) Profit vs. number of logins

Fig. 4. Trading success

success. A similar theme is noticed in Figure 4(a). The longer a user has been a member of the network, the less volatility they experience in the total profit/loss of their positions.

The above study makes a suggestion that the most influential nodes in Currensee might be found among those nodes with a small or moderate number of friendship links and log-ins, rather than among those with too many links or too many log-ins.

V. Conclusion

In our analysis of the friendship network of Currensee, we have found that user degree follows a power-law distribution. A power-law distribution also occurs for the network of private messages sent between users and the number of logins made by users. A disassortative mixing pattern is observed in the network. In terms of Forex trading behavior, we have found that the number of trades made by users in the network follows a power-law distribution.

These results are not particularly surprising. However, the users who trade the most have fewer number of friendship links within the network, as do the users who are most profitable in their trading. It seems likely that without some kind of incentive mechanism, the most active and most successful Forex traders remain reluctant to participate in the social aspects of an online social network tailored specifically for them. This raises the question: what kind of incentive is necessary here? If a monetary incentive is required, how much?

In our future work, it will be interesting to see whether these results will also hold as the network keeps growing. We will also investigate further into trading similarity by first determining a sufficient definition of what makes two different trades (or trading techniques) similar and then examining the tendency of users to form friendship links with other users who exhibit similar trading activity. It would also be interesting to measure the effect that the investing behavior of groups within the network has on the investing behavior of individuals.

Acknowledgment

The authors would like to thank the Currensee company for providing us with the dataset that was used for this study.

References

[1] P. Erdős and A. Rényi, "On random graphs," *Publicationes Mathematicae*, vol. 6, pp. 290–297, 1959.
[2] Currensee, http://www.currensee.com/.
[3] L. A. Adamic, O. Buyukkokten, and E. Adar, "A social network caught in the web," *First Monday*, 2003.
[4] R. Kumar, J. Novak, and A. Tomkins, "Structure and evolution of online social networks," *In Proc. ofKDD*, 2006.
[5] A. Mislove, K. P. Gummadi, and P. Druschel, "Exploiting social networks for internet search," *In Proc. of HotNets*, 2006.
[6] B. Viswanath, A. Mislove, M. Cha, and K. P. Gummadi, "On the evolution of user interaction in facebook," *Proceedings of the 2nd ACM SIGCOMM Workshop on Social Networks*, 2009.
[7] A. Mislove, H. S. Koppula, K. P. Gummadi, P. Druschel, and B. Bhattacharjee, "Growth of the flickr social network." *In Proc. of WOSN*, 2008.
[8] A. Mislove, M. Marcon, K. P. Gummadi, P. Druschel, and B. Bhattacharjee, "Measurement and analysis of online social networks," *Proceedings of the 5th ACM/USENIX Internet Measurement Conference*, 2007.
[9] A.-L. Barabasi and R. Albert, "Emergence of scaling in random networks," *Science, vol 286, no 5439*, 1999.
[10] H. Chun, H. Kwak, Y. Eom, Y. Ahn, S. Moon, and H. Jeong, "Comparison of online social relations in volumevs interaction: a case study of cyworld," *In Proc. of Internet Measurement Conference*, 2008.
[11] J. Leskovec and E. Horvitz, "Planetary-scale views on a large instant-messaging network," *In Proc. of the International World Wide Web Conference, Beijing, China*, 2008.
[12] G. P. Hopper, "What determines the exchange rate: economic factors or market sentiment?" *Business Review*, no. Sep, pp. 17–29, 1997. [Online]. Available: http://econpapers.repec.org/RePEc:fip:fedpbr:y:1997:i:sep:p:17-29
[13] A. Clauset, C. R. Shalizi, and M. E. J. Newman, "Power-law distributions in empirical data," vol. 51, no. 4, pp. 661–703, 2009. [Online]. Available: http://dx.doi.org/doi/10.1137/070710111
[14] M. E. J. Newman, "Mixing patterns in networks," *Physical Review E*, vol. 67, no. 2, pp. 026 126+, February 2003.
[15] ——, "Assortative mixing in networks," *Physical Review Letters*, vol. 89, no. 20, pp. 208 701+, October 2002.
[16] K. Nguyen and D. A. Tran, "An analysis of facebook activities," *In Proceedings of IEEE Consumer Communications and Networking Conference (CCNC 2011)*, 2009.