Characterizing the Activity of Friendship Triads on Facebook

Derek Doran, Alberto de la Rosa Algarin, and Swapna S. Gokhale

Abstract—Contemporary approaches to analyze user behavior on online social networks consider interactions among pairs of directly connected users or dyads. A large body of sociological work, however, suggests that mutual connections can influence the activity between two users, leading to differences between three-way and two-way interactions. This paper explores the dynamics of three-way relationships on Facebook. Categorizing each connection as a close friend or an acquaintance, contingent on the number of wall posts, this study examines how the different types of connections forming triads influence their characteristics embodied in posting activity, inter-post, reciprocation, and formation times, and activity decay. The analysis indicates that the number of constituent friendships primarily influences the characteristics of triads.

Index Terms—Behavioral analysis, facebook, online social network analysis, triads.

I. INTRODUCTION

Online Social Networks have captured our imagination by offering a revolutionary mechanism for communication and sharing. These massive networks hold rich troves of information that can be mined to understand patterns of social behavior in order to fulfill many objectives. For example, scientists may explore this behavior to assess if sociological theories established in the offline world transcend to online networks. Commercial organizations may also leverage these patterns to determine how users exert influence and recommend products and to run targeted advertising [1].

Longstanding sociological theories suggest that interactions differ fundamentally in a group of three as compared to a group of two people [2]–[4]. The relationship among such user triads has been studied extensively in the context of offline, physical social networks. Compared to a pairwise relationship, triadic relationship structures de-emphasize individuality, reduce power, and impose a natural moderation of conflict among the actors [5]. Triads de-value individuality because the opinion of one actor may be overwritten by the thoughts among the other two. Power is also lost compared to a dyad because in a dyad one actor can threaten to sever ties if some demands are not met, whereas in a triad such a threat is mitigated because a user who severs the tie is actually the one to become isolated from the other two actors. Finally, conflicts are more easily managed within a triad as the additional actor can mediate and resolve problems,

which may also reduce the likelihood of illegal or unethical behavior [3]. Triadic structures may also be unstable compared to dyads, as one of the three actors can quickly begin to feel unwanted or disconnected [2].

Most contemporary efforts that study online social networks consider dyadic relationships as the fundamental unit for analysis. Sociologists, however, consider "triads" or groups of three people users and the pairwise relationships among them as the fundamental unit of social network analysis because it is the smallest structural unit from which multi-person dynamics and interactions can be observed. Thus, computational social network analysis techniques that emphasize the properties of the activities that occur within triads rather than dyads are likely to unearth more sophisticated and insightful social patterns.

In this paper, we characterize the activity across different types of triads of Facebook users. We classify each pair-wise connection as a "close friendship" or an "acquaintance" based on the number of wall posts among the participating actors. Subsequently, we identify different types of triads based on the constituent pairwise close friend and acquaintance relationships. We then compare these different types of triads using four metrics, namely, their post volume, and inter-post, reciprocation, and formation times. Facebook wall posts, collected over approximately four years from the New Orleans regional network, provide the data for this study. Our analysis finds that the number of close friendships among the users influences these properties of triadic structures.

The paper is organized as follows: Section II describes how triads within our data set are classified. Section III compares the different types of triads using the various metrics. Related research is presented in Section IV. Conclusions and future work are offered in Section V.

II. CLASSIFICATION OF TRIADS

Formally, a triad is defined as three users whose connections form a complete graph. In our data set, we consider a connection between users P and Q if, at any time during the four-year period, P or Q posted a message on the other's wall. In this preliminary work, we do not consider triads containing structural holes [5], i.e., triples of users whose connections do not form a complete graph, or differentiate between triads composed of mutual dyads (e.g. a connection from P to Q and from Q to P).

Intuitively, it can be expected that not all Facebook connections are created equal. Some connections will feature a large volume of activity between two good friends, while others may show little activity between two acquaintances. Therefore, we classify each connection as a close friendship

or an acquaintance based on its strength, which we define in terms of the number of wall posts along the connection. To identify friendship and acquaintance connections, we first compute the mean number of posts along a connection, which for our data set are 3.24. We then designate a connection to be a close friendship only if the number of wall posts between the participating users exceeds this mean. Our definition of friendship requires strong activity only in one direction. In other words, if either P or Q exhibits behavior suggesting friendship (by posting at least 4 messages to the other's wall), we mark the relationship as a friendship. Furthermore, our friendship does not consider the cumulative number of wall posts along both directions (someone has to act like a friend for there to be a friendship). Thus, even if the total wall posts among P and Q exceeds the mean, P and Q are still acquaintances and not close friends if the number of posts in each direction is less than the mean.

Fig. 1 illustrates how this threshold splits the connections between friendships and acquaintances. Using the mean to split the connections is based on the following rationale. An "acquaintance" represents a very weak, and perhaps an even non-existent offline social tie between its two users. On the other hand, connections classified as "close friendships" have stronger ties among their participants as reflected by the higher number of wall posts among them. Splitting based on the mean labels approximately 80% of the connections as acquaintances. Given how users frivolously add connections to build social capital [6], we believe that this 80/20 split of connections into acquaintances and close friendships is reasonable.
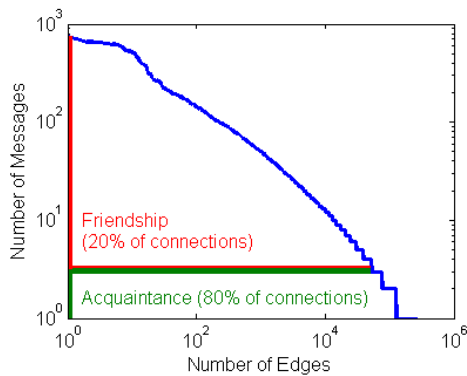

Fig. 1. Range of friendship connections

III. COMPARISON OF TRIADS

In this section, we quantitatively compare different types of triads using various metrics to characterize their properties. We also offer insights into the underlying social processes that influence these properties. For the purpose of this comparative analysis, we consider a data set of wall post activity off Facebook from the New Orleans regional network during from September 14th, 2004 through January 22nd, 2009 [7]. Table I offers a summary of the 876,933 wall posts from the 46,952 unique users in the data set. Approximately 46% of the wall posts are status messages, or posts by users to their own walls. The volume of activity, as measured in terms of the number of wall posts, is very low over a vast majority of connections; approximately 50% of all connections have just one wall post and the mean number of messages posted by a

user per day is only 3.24. Finally, the number of users who post messages to other's walls (40,981) is very similar to the number who receives messages on their walls (38,143).

TABLE I: SUMMARY OF WALL POST DATASET

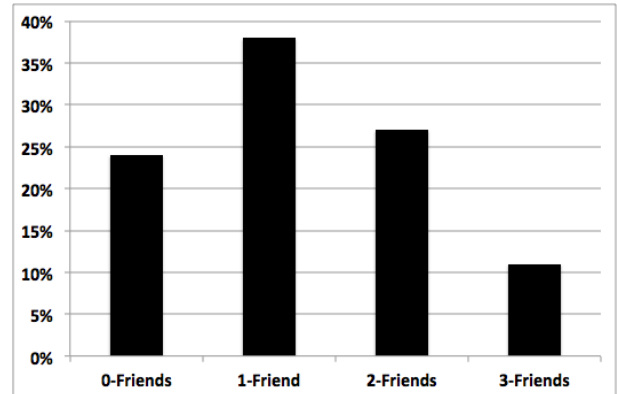|	
Total Posts	876,993
Num. of Users	46,952
Num. of Wall Posts	264,004
1 Message Connections	137,266
Status Messages	21,451
Mean Posts Sent	3.24

Fig. 2. Distribution of Triad Types

We generated a random sample of 100,000 triads from this dataset by collecting all triads that a user that is selected uniformly at random participates in. We repeated this process until we obtained 100,000 triads. Filtering out the duplicates, we were left with 89,526 triads. The classification of each connection defines four different types of triads based on the number of friendship edges they are composed of: *0-Friends*, *1-Friend*, *2-Friend*, and *3-Friends*. Fig. 2, which shows the distribution of the different types of triads, indicates that approximately one quarter of all triads consist of three acquaintance edges, supporting the notion that acquaintance connections on Facebook are added abundantly, without discretion [8]. The percentage of *1-Friend* or *2-Friends* triads is greater than the percentage of all acquaintances or all close friend's triads because while a full acquaintance/close friend triad only has one possible configuration, *1-Friend* and *2-Friends* triads each have three possible configurations. Finally, we observe that because a minority of Facebook connections is close friendships, the total percentage of triads decreases as the number of constituent close friendships increases. Next, we compare activity metrics for these different types of triads.

A. Posting Activity

Fig. 3 plots the reliability function on a log-log scale for the total number of wall posts in each type of triad. The trend for *0-Friends* triads is not shown because, by definition, the total activity across this type of triad cannot exceed 9 posts. If this were the case, the number of wall posts along at least one connection must exceed the threshold of 3 messages, which would then make it a close friend connection, and the triad would no longer be a *0-Friends* triad.
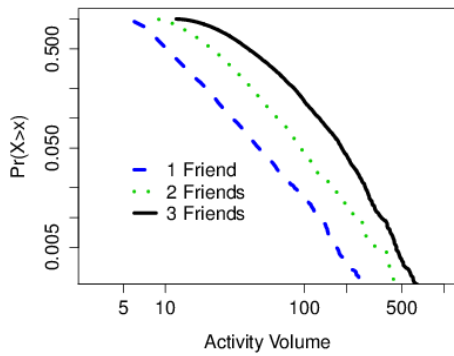
Fig. 3. Reliability Function of Post Volume

TABLE II: POWER-TAIL PARAMETERS OF POST VOLUME

Type	x_{min}	α	Mean
0-Friends	N/A	N/A	4.8
1-Friend	9	2.53	17.9
2-Frieds	37	2.68	35.6
3-Friends	174	3.45	63.7

The linear trend in the volume of posts across each type of triad on a log-log scale is a signature of a distribution that is power-tailed [9]. In a power-tailed distribution, the probability of witnessing a value greater than x is modeled as $R(x) \sim cx^{-\alpha}$ for $x > x_{min}$ where c is a constant, x_{min} is the value at which the power tail begins, and α is a parameter controlling the rate at which the probabilities of larger values decrease. Table II lists the estimates of x_{min} and α for each triad type using the maximum likelihood method [10]. It also summarizes the mean number of posts for the different types of triads. For triads with a single friendship, nearly the entire distribution is power-tailed with α = 2.53. When two close friendships are included, the power tail starts at higher values and the distribution exhibits less variation (α = 2.68). Finally, the power-tail for triads with three friendships starts significantly later and α = 3.45. Thus, x_{min} and α increase with the number of friendships, suggesting smaller variation and greater stability in the volume of activity across triads composed of a higher number of friends.
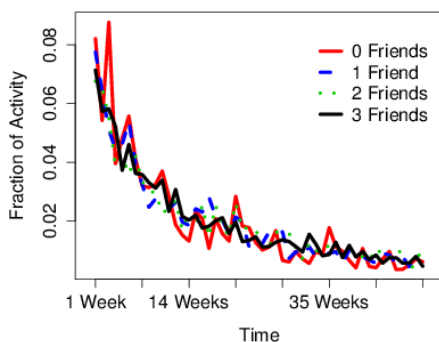

Fig. 4. Post Volume Decay Over Time

We also examine the rate at which post volume changes over time within the different triad types. Fig. 4 plots the fraction of total volume per week once a triad is formed. We exclude the first week because the volume in this week is disproportionately high across all types of triads. Subsequently, however, the post volume decreases sharply and converges to a low and similar rate for all triad types. This is surprising for triads with close friendships, because we

would expect the close friends to interact relatively frequently and uniformly. One possibility is that regardless of the triad type, the new edges and mutual connections cease to be novel after the first week. For example, the high initial activity may signal the group's recent event that triggered the formation of the friendship triad. As more time elapses since the event, however, the novelty of the friendship expires, leading to a decrease in activity.
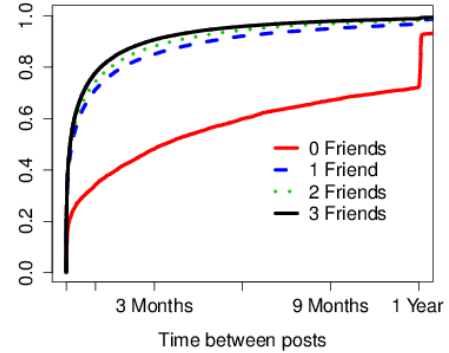

Fig. 5. Distribution of Inter-post Times

B. Inter-Post Times

We define inter-post time as the duration between successive posts along one direction of a connection of a triad. Fig. 5 plots the distribution of inter-post times measured across each pairwise connection for the different types of triads. Unlike post activity, which exhibited a remarkably different behavior depending on the number of close friendships involved, the distribution of inter-post times for *1-Friend* triads differs only moderately from the distribution for *2-Friends* triads. Table III, which presents a summary of the average time between posts in days for the different types of triads, confirms this observation. The table shows that as the number of close friendships in a triad increases, the mean inter-post time decreases. However, going from *1-Friend* to *2-Friends* triads, the average number of posts double (Table I), while the mean inter-post times reduce by only 25%. Similarly, going from *2-Friends* to *3-Friends* triads, the average number of posts increases by approximately 80% but the reduction in the mean inter-post time is only 20%. Thus, the shorter inter-post times along the one or more close friendships dominate the longer times along the complementary acquaintances. The sudden spike in the distribution for *0-Friends* triads at the one-year mark occurs because of the annual birthday messages commonly exchanged among Facebook users, which are also triggered and encouraged by Facebook alerts and reminders.

C. Reciprocation Times

We define reciprocation time as the difference between the time of a post from user P to Q, and the time of the subsequent first post from Q to P. Fig. 6 plots the distribution of reciprocation times across the different types of triads. Most messages are reciprocated within one month regardless of the type of triad. The distribution function increases slowly beyond the one-month mark. In other words, if a wall post on Facebook is not reciprocated within a short period of time, it is unlikely for a response to ever occur.

Table IV shows the mean reciprocation times of the different triads as a function of the number of friendships.

Although the reciprocation times decrease with the number of friendships similar to inter-post times, the progressive decrease in the percentage is comparatively lower. Specifically, going from *0-Friends* to *1-Friend* triads causes an 18% reduction, from *1-Friends* to *2-Friends* triads also results in 18% reduction, but moving from *2-Friends* to *3-Friends* triads results in only an 8% reduction.
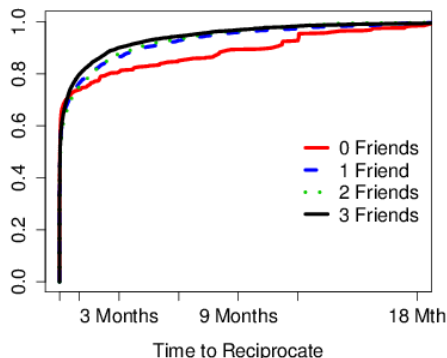

Fig. 6. Distribution of Reciprocation Times

TABLE III: MEAN INTER-POST TIMES

Type	Percentage
0-Friends	24 %
1-Friend	38 %
2-Fries	27 %
3-Friends	11 %

TABLE IV: MEAN RECIPROCATION TIMES

Type	Mean
0-Friends	66.7
1-Friend	44.39
2-Fries	36.9
3-Friends	35.2

D. Formation Times

We investigate the formation rates and times in order to examine whether the social relationships associated with triads of different types are reflected in these two metrics. We define the formation time of a triad as the earliest instant when at least one wall post was exchanged between a pair of users. Thus, the direction of the post is irrelevant for this definition. Fig. 7 plots the CDF of triad formation times. In this data set, the formation of triads is not significantly evident until August 2006. The rate at which the CDF increases reflects the rate at which user triangles form, which is a function of the number of constituent close friendships. The *3-Friends* triad forms at a nearly constant rate, which may be due to Facebook features that allow friends to find and connect with mutual friends easily. For example, users can browse the lists of their connections and also receive friend recommendations. As the number of close friendships in a triad decrease, the rate at which they form becomes sub-linear as the CDF takes a more concave shape. This may occur because acquaintances may be established more arbitrarily and on a whim on Facebook, whereas close friendships may be deliberately nurtured due to prior strong underlying social ties between the participating users in real life. However, as more impassive acquaintances are established, inevitably triads of unrelated users will form over time. This may further shrink the diameter of the social network [11], increasing the likelihood that a new connection would in turn form a user triangle.
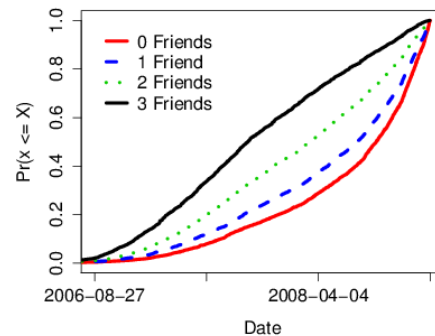

Fig. 7. CDF of Triad Formation Times

IV. RELATED WORK

Many efforts have examined user interactions in online social networks. Using the same data set, Viswanath *et al.* identified how interactions between dyads evolve [7]. Mislove explored the formation of new social ties on Flickr [8]. Cha investigated information flows on Twitter as users share and repost tweets [12]. Budak *et al.* consider pairwise interactions to limit the flow of misinformation [13].

Fewer efforts have considered activity among user triads. Romero *et al.* test the sociological theories of balance, exchange, and betweenness among triadic relationships on Twitter [14]. They also measures the directed closure processes in Triads on Twitter [15]. Meng *et al.* obtain theoretical results for the time taken to reach dynamic balance in triads [16]. Unlike these efforts, our study may be first to identify different types of triadic relationships on an online social network and empirically characterize their properties.

V. CONCLUSIONS AND FUTURE WORK

This paper identified different types of user triads on Facebook based on the strength of the social connections that bind them together. We compared these triads using several metrics that characterize their properties, namely, posting volume, inter-post, reciprocation and formation times, and activity decay. We found that the number of close friendships influences the properties of the triads.

Our future work seeks to further classify triads based on whether a friendship between two users is bi-directional. Investigating whether the properties of triads approach those of dyads as the number of close friendships increases is also a concern of the future. We will also examine the content of the information exchanged among users in different types of triads, to further determine whether offline sociological theories transcend to the online world.

REFERENCES

[1] M. Jamali, G. Haffari, and M. Ester, "Modeling the temporal dynamics of social rating networks using bidirectional effects of social relations and rating patterns", in *Proc. 20th Intl. World Wide Web Conference*, pp. 527–536, 2011.

[2] A. Baum, A. Shapiro, D. Murray, and M. V. Wideman, "Interpersonal mediation of perceived crowding and control in residential dyads and triads," *Journal of Applied Social Phychology*, vol. 9, no. 6, pp. 491–504, 1979.

[3] D. Brass, K. Butterfield, and B. Skaggs, "Relationships and unethical behavior: a social network perspective," *The Academy of Management Review*, vol. 23, no. 1, pp. 14–31, 1998.

[4] P. Holland and S. Leinhardt, "An omnibus test for social structure using triads," *Sociological Methods and Research*, vol. 7, pp. 227–256, 1978.

[5] G. Simmel and K. H. Wolff, *The Sociology of georg simmel*, Macmillan Publishing Co, 1950.

[6] N. Ellison, C. Steinfield, and C. Lampe, "The benefits of facebook 'friends:' social capital and college students' use of online social network sites," *Journal of Computer-Mediated Communication*, vol. 12, no. 4, pp.1143–1168, 2007.

[7] B. Viswanath, A. Mislove, M. Cha, and K. Gummadi, "On the evolution of user interaction in facebook," in *Proc. 2nd ACM Workshop on Online Social Networks*, 2009.

[8] A. Mislove, H. S. Koppula, K. Gummadi, P. Druschel, and B. Bhattacharjee, "Growth of the flickr social network,".in *Proc. of ACM SIGCOMM Workshop on Online Social Networks*, 2008.

[9] L. Lipsky, *Queueing Theory: A Linear Algebraic Approach,* Springer-Verlag, 2009.

[10] A. Clauset, C. R. Shalizi, and M. Newman, "Power-Law Distributions In Empirical Data," Technical report, arXiv:0706.1062v2, 2009.

[11] J. Leskovec, J. Kleinberg, and C. Faloutsos, "Graphs over time: densification laws, shrinking diameters, and possible explanations," in *Proc. 11th ACM SIGKDD Intl. Conference on Knowledge discovery in data mining*, pp.177–187, 2005.

[12] M. Cha, A. Mislove, and K. Gummadi, "A measurement-driven analysis of information propagation in the flickr social network," in *Proc. 18th Intl. Conference on World Wide Web,* 2009.

[13] C. Budak, D. Agrawal, and A. E. Abbadi, "Limiting the spread of misinformation in social networks," in *Proc. of 20th ACM Conference on the World Wide Web*, 2011.

[14] D. Romero, B. Meeder, V. Barash, and J. Kleinberg, "Maintaining ties on social media sites: The competing effects of balance, exchange, and betweenness," in *Proc. 5th International AAAI Conference on Weblogs and Social Media*, 2011.

[15] D. Romero and J. Kleinberg, "The directed closure process in hybrid social-information networks, with an analysis of link formation on twitter," in *Proc. 4th Intl. AAAI Conference on Weblogs and Social Media*, 2010.

[16] Q. K. Meng, W. Liu, and J. Y. Zhu, "Ternary Social Networks: Dynamic Balance and Self-Organized Criticality," Technical report, 2010.

Derek Doran is a Ph.D. candidate in the Department of Computer Science and Engineering at the University of Connecticut. His broad research interests are in big data analytics, social network analysis, and systems performance modeling. His current research focuses on understanding the performance implications of Web crawler traffic on the Internet, large-scale network analysis, and social media mining with applications to the power utility, transportation system, telecommunications, and healthcare industries. Derek is a National Science Foundation EAPSI Fellow, a Transportation Research Board Graduate Research Award recipient, a twice Summer Alumni of Bell Labs in Murray Hill, NJ, and a Best Paper award winner at the 2012 IEEE Intl. Conference on Machine Learning and Applications. He is a student member of ACM, IEEE, SIAM, and the TRB.

Alberto de la Rosa Algarin a third year Ph.D. student of Computer Science & Engineering at the University of Connecticut, with Dr. Steven A. Demurjian as major advisor. His research interests include: security and information assurance by utilizing document-level security, privacy enforcement, security policy integration and automatic generation, security and knowledge modeling, secure information and data engineering, identity-inferred access control. His current research focuses on UML meta-model extensions for XML security and policy integration, with the end purpose of defining a software engineering process for integrated information exchange meta-systems. He received a B.S., with a major in Computer Science and a second major in Mathematics, from the University of Puerto Rico, where he had research experiences involving software engineering and reliable distributed systems.

Swapna S. Gokhale is an associate professor in the Dept. of Computer Science and Engineering at the University of Connecticut, Storrs, CT, USA. She received her B.E. (Hons.) in Electrical and Electronics Engineering and Computer Science from the Birla Institute of Technology and Science, Pilani, India in 1994 and her M.S. and Ph.D. in Electrical and Computer Engineering from Duke University, Durham, NC, USA in 1996 and 1998 respectively. Her research interests include mining social media data, software engineering education, and performance and dependability analysis of computer systems. She is a Senior Member of the IEEE, recipient of the NSF CAREER Award, and best paper awards at several international conferences.