The Workshops of the Thirtieth AAAI Conference on Artificial Intelligence
Incentives and Trust in Electronic Communities: Technical Report WS-16-09

# An Analysis of Trimming in Digital Social Networks

**Renita Murimi**
Oklahoma Baptist University
renita.murimi@okbu.edu

**Abstract**

The study of network sizes in digital social networks is a research question of significant interest. Here, we explore the phenomenon of trimming, which is the decrease in the size of one's network, and analyze if the rules of social exchange theory – namely, status consistency and reciprocity- can affect trimming. To this end, we use a Hidden Markov Model to investigate the relationship between the frequency of interaction and one's network size, in which we are able to control for the current size of one's digital social network. We find that there are significant patterns in sharing tendencies in digital social networks. One is that users who do not share enough are the group that is most likely to be trimmed from a network. Another is that users prefer to have moderate sized networks, i.e. networks with 500 – 1000 friends and prefer friends with moderate sharing tendencies (sharing approximately once a week). We also find that one's sharing preferences over time tend to align with moderate sharing.

*Keywords: Digital social networks, Hidden Markov Models, Social Exchange Theory*

## Introduction

Digital social networks (DSNs) are complex social networks (Barabasi 2002; Newman 2003) that have created *friends* and *followers, pins* and *tags*. A little more than a decade ago, these terms might not have meant anything more than their conventional meaning implied by companion, adherent, tack and label respectively. However, DSNs such as Facebook, Twitter and Pinterest have turned these terms into cornerstones of the digital social media lexicon. According to a January 2015 report at wearesocial.net, roughly 29% of the world population (approximately 2 billion people) have active user accounts on social media. Facebook is the most popular social media platform with about 1.3 billion users. Users also spend a significant amount of time on social networks: the average amount of time spent by users on Facebook is now around 20 minutes a day. On August 24, 2015, Facebook reached a critical milestone: for the first time ever, one billion people used Facebook in a single day. How are users of DSNs managing their friends, followers and tags?

In this paper, we focus on user behavior in Facebook as a function of network size and frequency of interaction. Facebook is a social media platform that enables a user to find contacts – *friends*– and create networks. Friends of friends are automatically created when a friend expands his/her network, thereby creating the potential to form deep networks of hundreds of thousands of connections. Users have the ability to post content in the form of opinions, links, and multimedia and also to join or form specific networks, for e.g. the Farmer's Market in Shawnee, OK. A user's activity appears on the news feeds of friends in his/her network. Facebook defines news feed as the constantly updating list of stories in the middle of one's home page that includes status updates, photos, videos, links, app activity and likes from people, Pages and groups that one follows on Facebook.

The ability to create networks of friends has brought with it interesting challenges. Since friendship in a DSN such as Facebook, is merely a function of a button click that reads "Send friend request", casual acquaintances, contacts from a business dinner, distant relatives and even people one has never met in person can all become *friends* on Facebook. In (Gladwell 2000), the author presents the number of 150 friends as the optimal size for a community based on Dunbar's study (Dunbar 1992) on the number of meaningful stable relationships that one can form. In a separate study, college students examined (Tong et al. 2008) the number of friends in various networks. The study showed that a network of around 300 friends was the most appealing.

Users manifest their desire to form smaller, more meaningful networks in various ways. Clicking on familiar icons and links enables the formation of a friendship, and options to *unfriend/unfollow/hide* can temporarily or permanently disable the friendship. Multiple such ways of friendship control (and hence, network size control) exist: on one end of the spectrum, this can be achieved using manipulation of privacy settings. Privacy settings enable users to control the access of their information by friends on their network, thereby letting only a subset of the total number of friends to see their posts. Multiple such subsets can be created with varying levels of access. On the other end of the spectrum, users can *unfollow* a friend, choose not to see posts from the friend or even delete a friend, following which that friend is no longer a part of the user's feed or network. What causes users to reduce the size of their networks? In digital networks containing hundreds or thousands of friends, users might form informal sub-networks of friends with whom they interact more frequently than the others. In this paper, we define and study *trimming* as the phenomenon of reducing the size of one's DSN, either through pre-defined formal mechanisms (such as *unfriend*, *unfollow*, privacy settings or *hide posts*) or through informal situational contexts such as regularly communicating only with a certain subset of friends on Facebook. We focus on the impact of posted content on trimming, i.e., could the frequency of one's posts cause his/her friends to reevaluate the need for friendship with that person on Facebook? In other words, can sharing a lot or not sharing enough get you booted from someone's network? A person's comments, pictures, likes, shares and status updates may be viewed as positive or offensive, irrelevant or uninteresting. With one's status in a DSN at stake through the frequency and content of the posts, are there certain thresholds (size of network, content of a post, frequency of posts) that need to be adhered to ensure a successful existence on DSNs? We consider, whether and if, there exist rules in social exchange that members of a DSN could adopt to ensure optimal behavior in a DSN.

## Related Work

Various DSNs have been studied through the tools available in computational social science, machine learning and computational linguistics. In (Danescu-Niculescu-Mizil, Sudhof, and Jurafsky 2013), the authors study the linguistic aspects of politeness by constructing ac corpus of requests made to Wikipedia editors, where they found that polite Wikipedia editors achieve high status, but once elevated in status, they become less polite. The problem of discovering social circles in Google+, Facebook and Twitter has been studied in (Leskovec and Mcauley 2012), where the problem of discovering social circles is posed as a node clustering problem that is analyzed with machine learning algorithms. Network formation in research collaborations has been explored in (Dahlander and McFarland 2013) to analyze the impact of shared interest on the formation and sustenance of collaborations. Patterns of information sharing have been studied in online networks (Rafaeli and Raban 2005) and in organizations (Constant, Kiesler and Sproull 1994). Another study to infer human behavior from their DSN data has been performed in (Barchiesi et al. 2015), where the authors analyzed photographic data shared on Flickr to infer human mobility patterns to find the probability of people in geographic locations and the probability of movement between locations. Multiple other studies have investigated the factors that influence human behavior in various kinds of social networks. In (Fowler, Dawes and Christakis 2009), the authors study the impact of genetic variation on human behavior in social networks. The results show that the genetic variation affect how many times a person is named as a friend, but does not significantly impact how many friends a person names. The posts on Twitter (tweets) have been studied to infer situational information (Saleem, Xu, and Ruths 2014), exercising behavior (Jurgens, McCorriston, and Ruths 2015), political orientation (Cohen and Ruths 2013) and gender information (Ciot, Sonderegger, and Ruths 2013). Our work addresses the relationship between frequency of sharing and network sizes, and we draw upon approaches from social exchange theory (Blau 1960; Meeker 1978; Cropanzano 2005) and Hidden Markov Models to develop a framework to explore this further.

## Methodology

In this paper, we study the factors that cause trimming of networks. Specifically, we investigate the relationship between frequency of sharing by one's friends and the size of one's network. In order to study the relationship between network size, frequency of sharing and trimming, we use the framework developed in (Meeker 1978), where social interactions are framed as belonging to a set of exchange rules in the categories of rationality, altruism, status consistency or competition. The work in (Meeker 1978) develops this framework for normative rules of social exchange using the rules of game theory, where rewards and costs are associated with people's behavior in social exchange. Of these categories, we study how status consistency and reciprocity influence user behavior in DSNs. Game theory is a theory of decision making under conditions of uncertainty and interdependence. We now provide a brief introduction of the basics of game theory. A game has three components: a set of players, a set of possible actions for each player and a set of strategies. A player's strategy is a complete plan of actions to be taken when the game is actually played. Players can act selfishly to maximize their gains and hence a distributed strategy for players can provide an optimized solution to the game. In any game, utility represents the motivation of players. A utility function, describing player's preferences assigns a number for every possible outcome of the game with the property that a higher number implies that the outcome is more preferred. For two individuals A and B involved in a social exchange, (Meeker 1978) defines status consistency as an exchange rule that assigns the maximum value to the difference between the payoffs of A and B. On the other hand, reciprocity has been defined as an exchange rule that assigns the minimum value to the difference between the amount that A's decisions have contributed to B's payoff and the amount that B's decisions have contributed to A's payoff.

The following terms will be introduced and defined in the context of DSNs for our study of trimming:

- Frequent sharers: Users whose frequency of posting on Facebook exceeds once per day.
- Moderate sharers: Users whose frequency of posting on Facebook is at least once per week, but less than that of frequent sharers.
- Sparse sharers: Users whose frequency of posting on Facebook is at least once per month, but less than that of frequent sharers and moderate sharers.
- Non-sharers: Users who you know are on your network, but you do not see any updates from them.
- Post: Status updates, pictures, links, shares, likes.
- Small network: 1- 500 friends.
- Medium network: 500 -1000 friends.
- Large network: Greater than 1000 friends.

Where do DSNs fit into the framework of social exchange theory? Rewards are manifested in approval, which in Facebook is implemented using *likes*, *shares* and positive comments. The costs are manifested in the form of negative comments and the mechanism of *unfriending*, *unfollow*ing or having one's activities being hidden from one's friends' feeds. We use a HMM to model the time series of a user's sharing patterns (frequent, moderate, sparse or non-sharers) and investigate the effect of network sizes on sharing preferences.

## Hidden Markov Model Analysis of Sharing Patterns

A Hidden Markov Model (HMM) is a kind of dynamic Bayesian network that can be used to represent probability distributions over observation sequences, and thus are used for modeling time series data (Ghahramani 2001). HMMs represent conditional probabilities of dependence between random variables. In order to use HMMs to learn about trimming in networks, initial knowledge about the sharing preferences and network size are applied to obtain a posterior distribution. We use this kind of predictive distribution afforded by HMMs to predict the most likely sharing preferences of users over time (Stamp 2004).

Initial knowledge: We model a person's sharing preferences as belonging to one of four categories: frequent sharers (F), moderate sharers (M), sparse sharers (S) and non-sharers (N). These four categories are the states of a Hidden Markov Model, where the transition from one state to another is a Markov process of order one. Further, a user's network is modeled as comprising of the above four categories of sharers. The matrix in equation (1) indicates the probability of an individual being friends with the same or another category of sharers. For example, the first row of this matrix indicates that a frequent sharer has a probability that $a_1$ % of the friends in his/her network are frequent sharers, a probability that $a_2$ % of the friends in his/her network are moderate sharers, a probability that $a_3$ % of the friends in his/her network are sparse sharers, and a probability that $a_4$ % of the friends in his/her network are non-sharers. Similar statements can be made for the probabilities $b_i, c_i$ and $d_i$, where $i \in \{1,2,3,4\}$. The transition from one state to another can be effected by trimming the network or by expanding the network. For example, the probability $a_i$ can be increased by trimming the number of friends in other categories or by adding more friends in the category $a_i$. The state transition matrix $X$ denoting these probabilities is as follows:

$$X = \begin{array}{c c} & \begin{array}{cccc} F & M & S & N \end{array} \\ \begin{array}{c} F \\ M \\ S \\ N \end{array} & \begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \\ d_1 & d_2 & d_3 & d_4 \end{bmatrix} \end{array} \qquad (1)$$

The relationship between the size of a person's network (small, medium or large) and the category of sharing (frequent, moderate, sparse, non-sharers) is summarized in a row stochastic matrix $Y$ in equation 2 as a probabilistic relationship between the two entities. The first row denotes the probability that a frequent sharer has a small network $(S_n)$ is $x_1$, has a medium network $(M_n)$ is $x_2$ and that the probability of a large network $(M_n)$ is $x_3$. Similar statements can be made for the probabilities $y_i, z_i$ and $u_i$, where $i \in \{1,2,3\}$. The observation matrix $Y$ is as follows:

$$Y = \begin{array}{c c} & \begin{array}{ccc} S_n & M_n & L_n \end{array} \\ \begin{array}{c} F \\ M \\ S \\ N \end{array} & \begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \\ u_1 & u_2 & u_3 \end{bmatrix} \end{array} \qquad (2)$$

The initial state distribution, denoted by $\pi$ is given as a row stochastic matrix,

$$\pi = [p_1 \quad p_2 \quad p_3 \quad p_4] \qquad (3)$$

where the probabilities $p_i$ denote the initial probability distribution of the categories $F, M, S$ and $N$.

The question we are trying to answer is this: For a social network user, what is the most likely pattern of sharing as a function of time? In other words, we investigate the relationship between network size and frequency of sharing. For example, does a frequent sharer exhibit a change in sharing patterns over time as a result of change in his/her network size? Mathematically speaking, this research aims to identify the most likely state sequence of the Markov process given in equation (1).

| Observation sequence of network sizes | Most likely sharing tendencies | Least likely sharing tendencies |
|---|---|---|
| {S,M,S,M, L} | {Sp, Mo, Mo, Mo, Mo} | {Fr, No, Mo, Fr, No} |
| {S,M,L,M,S} | {Sp, Mo, Mo, Mo, Mo} | {Fr, No, No, Fr, No} |
| {S,M,L,M,L} | {Sp, Mo, Mo, Mo, Mo} | {Fr, No, No, Fr, No} |
| {S,M,S,M,S} | {Sp, Mo, Mo, Mo, Mo} | {Fr, No, Mo, Fr, No} |

*Table 1: Most likely sharing tendencies. Legend: Fr: Frequent, Mo: Moderate, Sparse: Sp, No: Non-sharers*

Our HMM model has 4 states given by ( $F, M, S$ and $N$ ) and 3 observation symbols ( $S_n$, $M_n$ and $L_n$ ). We choose an observation sequence of five sizes in order to observe the possible trimming and expansion of networks over a period of time. Hence, we consider a generic state sequence of length five:

$$Q = ( q_0 \quad q_1 \quad q_2 \quad q_3 \quad q_4 ) \tag{4}$$

with corresponding observations

$$O = ( o_0 \quad o_1 \quad o_2 \quad o_3 \quad o_4 )$$

(5)

Thus $\pi_{q_0}$ is the probability of starting in state $q_0$ . Consider the matrix $A = \{a_{ij}\}$ . This matrix is $n \times n$ with

$$a_{ij} = P(state \;\; j \text{ at } t+1 \,|\, \text{state } i \text{ at } t)$$

Similarly, the matrix $B = \{b_j(k)\}$ is a $n \times m$ matrix with

$$b_j(k) = P(observation \;\; k \text{ at } t \,|\, \text{state } j \text{ at } (t))$$

Hence the probability of the state sequence $Q$ is given by

$$P(Q) = \pi_{qo} * b_{q_0}(O_0) * a_{q_0,q_1} * b_{q_1}(O_1) * a_{q_1,q_2} * b_{q_2}(O_2) * a_{q_2,q_3} * b_{q_3}(O_3) * a_{q_3,q_4} * b_{q_4}(O_4) \tag{6}$$

For our model of four states and an observation sequence of length five, the number of possible state sequences are $4^5$ . The optimal sequence of most likely sharing tendencies in the HMM model is obtained by summing the probabilities that have a state in a given position and choosing the state with the highest normalized probability.

## Experimental Setup

A sample size of 118 students was used to obtain the data in this study. Students at Oklahoma Baptist University were administered a paper-and-pencil based survey. We use the results of our survey to populate a 4 x 4 row stochastic matrix $X$ , given in equation 1.

We consider an observation sequence for equation (5) of gradually varying states, where a network can transition in size only to the next lower or upper size. So, an observa tion sequence of five sizes: {S,M, S,M, L} is a valid sequence whereas the sequence {S, L, M, S, M} is an invalid

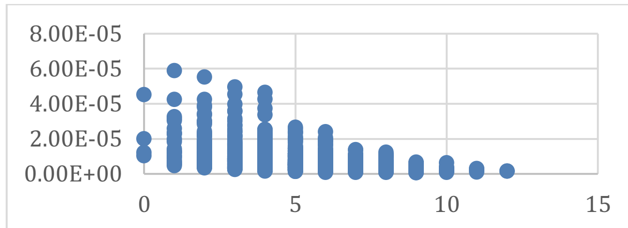


*Figure 1: Probability of sharing patterns as a function of change in network sizes over a period of time for a given observation sequence.*

observation sequence due to the transition from a network of small size to one of a large size. Thus, the four valid observation sequences are {S,M,L,M,S}, {S,M,S,M,L}, {S,M,L,M,L} and {S,M,S,M,S}. For these four valid observation sequences, we determine the most likely sequence of states of a person's network. That is, we can obtain the most likely probability of a user's sharing tendencies over the period of the observation sequence. Table 1 lists the most likely and least likely pattern of sharing tendencies as a function of each of the four valid observation sequences. Additionally, we introduce a variable $n$ that describes the differences in sharing patterns. We assigned the values 0, 1, 2 and 3 to the different categories of sharing of users (frequent, moderate, sparse and non-sharers) in order, respectively. We define $n$ as the sum of the absolute differences between consecutive states. Thus,

$$n = \sum_{i=0}^{4} |o_i - o_{i+1}|$$

Figure 1 shows a scatter plot graph of the probability of sharing patterns as a function of $n$ for 1024 possible combinations of sharing preferences over five time periods of observation sequences. We see that the higher values of $n$ correspond to sudden changes in sharing patterns. A gradual change in sharing patterns would transition between sharing frequencies smoothly, for example: {frequent, moderate, sparse, moderate, sparse}. Figure 1 shows that sharp transitions (higher $n$ values) are less common than gradual transitions. The most likely combination found by our HMM model has an $n$ value of 1 and corresponds to the pattern given by {sparse, moderate, moderate, moderate, moderate}. These results suggest that the social exchanges that occur in DSNs follow the set of exchange rules found in regular human networks and tend to align with the exchange rules of social exchange theory for status consistency and reciprocity. Users evaluate the nature of their interactions in terms of frequency and content of posts on Facebook and consider the impact it might have on the friends in their network. Posts that might be deemed as negative to a friendship have the effect of leading to network trimming.

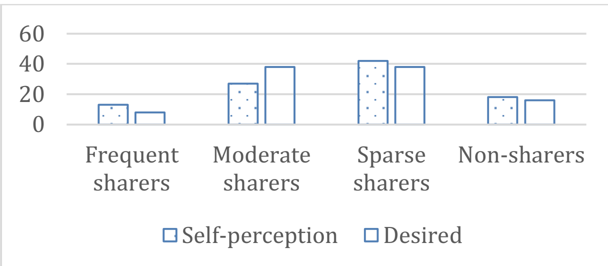


*Figure 2: Self-perception and desired level of sharing.*

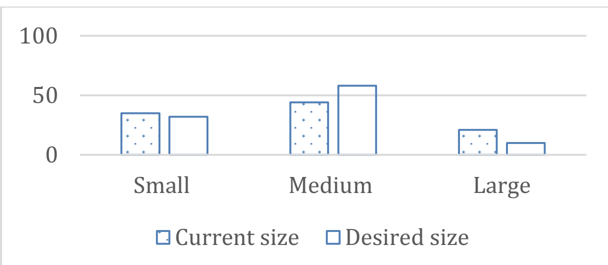


*Figure 3: Current and desired network sizes.*

## Results

### Sharing patterns and preferences

The results of our survey are shown in Figures 2- 7. Based on the definition of the various kinds of sharers, students were asked to place themselves in a particular category of sharing (self-perception) as well as choose their desired category of sharing (desired). Figure 2 shows the percentages of students in various categories. The category with the highest number of sharers is sparse sharers (49%). The category with the lowest number of sharers is that of frequent sharers (15%). For desired sharing levels, Figure 2 shows that sparse sharing and moderate sharing (45% each) were the top choices, compared to frequent sharing (9%) which was the least desired level of sharing. These results of the self-perception and desired level of sharing show that sharing on DSNs follows a pattern where the frequency of posts tend to stay in the sparse category (posts occur approximately once a month).

### Network size distribution

Figures 3 and 4 show the network size distribution of students. Figure 3 shows that 52% of the networks were of medium size and comprise the largest group. The smallest group was of large networks (25%). DSNs such as Facebook offer convenient tool to increase the size of a network by “suggesting friends” or “sending friend requests”. These mechanisms enable fast proliferation of network size, and often users have large network sizes.

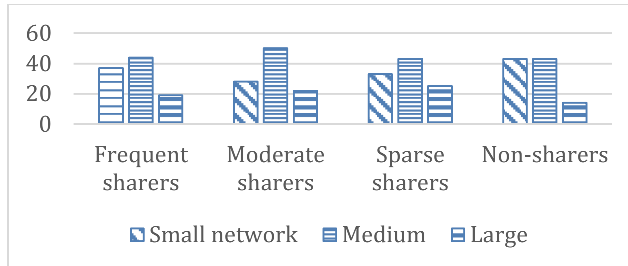


*Figure 4: Network sizes of different types of sharers.*

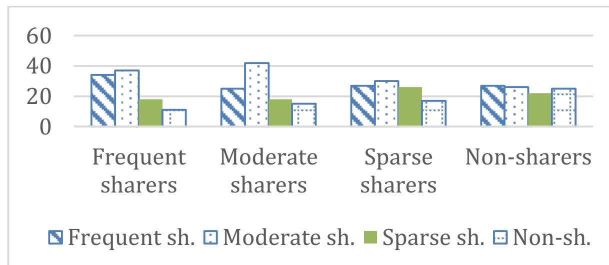


*Figure 5: Perceptions of sharing*

We asked students to also choose their preferred network size. Figure 3 shows the most preferred network size is a medium sized network (500 – 1000 contacts). The least preferred network size was a large network (12%). Thus, our surveyed population shows that among people who have large networks, the network size among half of them was something that they did not prefer to have. The results highlight the nature of digital friendships – sometimes our DSNs might contain friends that we would rather not have on that network. Figure 4 depicts the network size distribution of different types of sharers. Large networks were the least likely size across categories of sharers. Medium sized-networks were the most prevalent network size for frequent, moderate and sparse sharers. For non-sharers, small and medium sized networks were equally most likely.

### Impact of sharing on trimming

In Figure 5, students were asked to assign a percentage to the number of friends they would deem as belonging to a particular sharing category. The results were distributed across the self-perception of sharing. For example, frequent sharers reported that 34% of their friends were frequent sharers. Survey results show that across three categories of sharers (frequent, moderate and sparse), the most likely category of friends were moderate sharers. Non-sharers, however, thought that most of their friends were frequent sharers, followed very closely by moderate sharers. The same pattern of perception holds for the least likely group of sharers. Across three categories of sharers (frequent, moderate and sparse), the least likely category of friends were non-sharers.

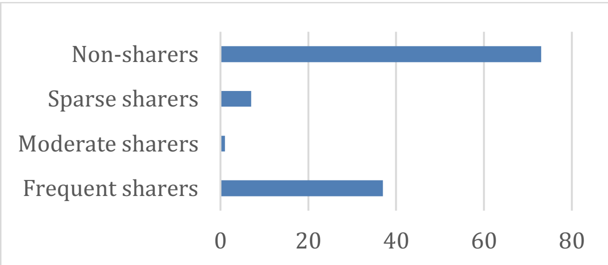


*Figure 6: Role of sharing in trimming*

Non-sharers, however, thought that the least likely sharing preference of their friends would fall into the sparse category. Figure 6 describes the impact of sharing on trimming. Students were asked to choose the category of friends they would trim based solely on the frequency of sharing. The results show that non-sharers would be trimmed by 73% of the population. It also shows that moderate sharers are the group that were least likely (1%) to be trimmed. We further studied the impact of perceptions of sharing frequency and distributed those across one's self-perception of sharing. Figure 7 shows the results of this analysis. Non-sharers were the group that were most likely to be trimmed across all categories of sharers. Along the lines of Figure 6, moderate sharers were the least likely group to be trimmed.

## Conclusions

In this paper, we studied the impact of frequency of sharing and network size on trimming. The results of our survey pointed out three salient observations: (1) Users tend to prefer medium-sized networks. (2) Moderate sharers are the group that is most preferred in a network for their sharing tendencies. (3) If the users of a DSN such as Facebook were to trim their network solely based on the frequency of sharing, non-sharers were the group that are most likely to be trimmed. The last observation also holds implications for businesses that use Facebook for marketing. Non-sharers do not add significant content, and hence are not much valued among a business's network of *followers.* We provide three primary factors that are likely for this final observation of trimming non-sharers over other groups. (a) Non-sharers are not interested in maintaining an active presence on DSNs and therefore do not login and update their posts, or (b) non-sharers tend to 'snoop' by just observing the activity of friends on their network and not updating their profiles or (c) the non-sharer has blocked a particular friend from viewing their posts on his/her feed. In so doing, it may appear that the non-sharer is not active on the network but in reality, the non-sharer has chosen to trim their network formally or informally thereby leading a friend to believe that he/she is a non-sharer. Finally, we used a Hidden Markov Model to determine that users'

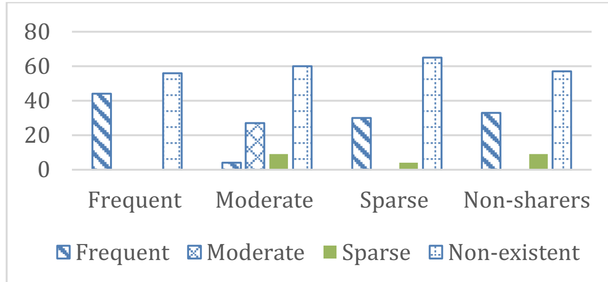


*Figure 7: Trimming distribution as a function of sharing.*

most likely sharing patterns tend to be of moderate frequency. These results were also verified using the metric of change in sharing patterns. The most probable sharing patterns exhibit a low value of $n$. In other words, people do not like rapid changes in the frequency of sharing. Our results show that user behavior in DSNs in terms of sharing align with exchange rules of reciprocity and status consistency. Users tend to form networks with others with whom they share similar contexts. In this process, the rewards and costs of DSN posts are evaluated to avoid negative outcomes like trimming.

## Future Work

While we explored the relationship between sharing and trimming for this paper, there are a number of questions that merit attention for further investigation. How do people exhibit rationality in their transactions on DSNs? Are 'likes' and 'shares' motivated not just by interest in the topic, but also by reciprocity? Does a person's affiliation with certain groups increase their status in their network, thereby triggering an increase in the number of likes? Or does competition for approval trigger actions of a certain kind among people who have 'mutual friends'? Most of all, is digital friendship prone to the same normative rules of social exchange as friendship in real-world settings? The research potential at the intersection of computational social science and DSNs are many due to the seamless interaction between people and their many digital social networks. The insights gleaned from such analyses will be important as we shape the discourse in modern computational social science for DSNs.

## References


Barabasi, A.-L. 2002. Linked: The New Science of Networks. Basic Books.

Blau, P.M. (1960) Structural Effects. American Sociological Review. 25(2): 171- 193.

Barchiesi, D.; Preis, T.; Bishop, S.; and Moat, H.S. 2015. Modelling Human Mobility Patterns using Photographic Data Shared Online. *Royal Society Open Science* 2 150046.

Ciot, M.; Sonderegger, M.; and Ruths, D. 2013. Gender Inference on Twitter in Non-English Contexts. In Proceedings of the Conference on Empirical Methods on Natural Language Processing.

Meeker, B. F. 1978. Decisions and Exchange. *American Sociological Review.* 485-495.

Cohen, R.; and Ruths, D. 2013. Political Orientation Inference on Twitter: It's Not Easy! In Proceedings of the International AAAI Conference on Web and Social Media.

Constant, D.; Kiesler, S.; and Sproull, L. 1994. What's Mine is Ours, Or Is It? A Study of Attitudes about Information Sharing. *Information Systems Research* 5: 400 – 421.

Cropanzano, R. 2005. Social Exchange Theory: An Interdisciplinary Review. *Journal of Management*. 31: 874-900.

Dahlander, L.; and McFarland, D.A. 2013. Ties that Last: Formation and Persistence in Research Collaborations over Time. *Administrative Science Quarterly* 58(1): 69 – 110.

Danescu-Niculescu-Mizil, D.; Sudhof, M.; and Jurafsky, D. 2013. Computational Approach to Politeness with Social Factors. In Proceedings of the Association for Computational Linguistics.

Dunbar, R.I.M. 1992. Neocortex Size as a Constraint on Group Size in Primates. Journal of Human Evolution 22(6): 469-493.

Fowler, J. H.; Dawes, C. T.; and Christakis, N.A. Model of Genetic Variation in Human Social Networks. In Proceedings of the National Academy of Sciences of the United States of America 106.6 (2009): 1720–1724.

Ghahramani, Z. 2001. An Introduction to Hidden Markov Models and Bayesian networks. International Journal of Pattern Recognition and Artificial Intelligence. 15(1): 9 – 42.

Gladwell, M. 2000. The Tipping Point: How Little Things Can Make a Big Difference. Boston: Little, Brown, and Company.

Gouldner, A. 1960. The Norm of Reciprocity. American Sociological Review. 25: 161-178.

Jurgens, D.; McCorriston, J.; and Ruths, D. 2015. An Analysis of Exercising Behavior in Online Populations. In Proceedings of the International AAAI Conference on Web and Social Media.

Leskovec, J.; and Mcauley, J. 2012. Learning to Discover Social Circles in Ego Networks. In Proceedings of the Conference on Advances in Neural Information Processing Systems.

Newman, M.E.J. 2003. The Structure and Function of Complex Networks. SIAM Review 45: 167-256.

Rafaeli, S.; and Raban, D.R. 2005. Information Sharing Online: A Research Challenge. *International Journal of Knowledge and Learning* 1(1-2), 62-79.

Saleem, H.M.; Xu, Y.; and Ruths, D. 2014. Novel Situational Information in Mass Emergencies: What does Twitter Provide? In Proceedings of the Conference on Humanitarian Technology: Science, Systems and Global Impact.

Stamp, M. 2004. A Revealing Introduction to Hidden Markov Models, Technical Report, Department of Computer Science, San Jose State University.

Tong, S. T.; Van Der Heide, B.; Langwell, L.; and Walther, J.B. 2008. Too Much of a Good Thing? The Relationship between Number of Friends and Interpersonal Impressions on Facebook. Journal of Computer-Mediated Communication 13(3): 531-549.