Measuring Behavioral Trust in Social Networks

Sibel Adali, Robert Escriva, Mark K. Goldberg, Mykola Hayvanovych, Malik Magdon-Ismail,
Boleslaw K. Szymanski, William A. Wallace and Gregory Williams

***Abstract*—Trust is an important yet complex and little understood aspect of the dyadic relationship between two entities. Trust plays an important role in the formation of coalitions in social networks and in determining how high value of information flows through the network. We present algorithmically quantifiable measures of trust based on communication behavior. We propose that trust results in likely communication behaviors which are statistically different from random communications; detecting these trust-like behaviors allows us to develop a quantitative measure of who trusts whom in the network. We develop algorithms to efficiently compute such *behavioral trust* and validate these measures on the Twitter network.**

I. Introduction

Trust is an important aspect of the relationship between two entities. The trust landscape of a social network (who trusts whom) plays an important role in the intelligence and security domain. Trust forms a basis for formation of coalitions (strong communities are formed by entities which "trust" each other); it can serve to identify influential nodes in a network; and, it determines how information will flow in a social network. The reverse is also true: communities can induce greater trust among the members; continued information flow between members can enhance the trust relationship between them.

Trust is a complex relationship influenced by a host of factors, such as: 1) Our own predisposition to trust which itself was influenced by various events over our lifetime; 2) Our relationship and past experiences with the person and with her friends. 3) Our opinions of actions and decisions the person has made. 4) Rumors and opinions of other people, and our trust in those people. Thus, trust in social networks is not yet well understood. To quantify trust, we must focus on some specific properties of trust, which may have to be simplified, so that these properties may be captured algorithmically.

This paper is about quantitatively measuring dyadic trust (trust between two entities) based on observed communication behaviors – we call this *behavioral trust*. A typical social network consists of actors (individuals) and communications between them (phone calls, emails, blog posts, etc). OUr challenge is to quantify trust only on the basis of the observed communication behavior (a portion of the interactions between entities). A useful analogy is "imitation is the best form of flattery" – imitation is a behavior indicative of a certain relationship, and similarly there are behaviors indicative of trust.

Our contribution is the development of algorithmically quantifiable measures of trust. The basis for our study is the proposition that trust results in a number of likely patterns of behavior. By measuring these patterns, we measure behavioral trust. Our measures are *statistically* defined, and do not use any semantic information in the messages; our algorithms are efficient and scale to million-node networks.

Related Work. There is work on trust in computer science as well as in social science. The topics range from: formation, emergence, valuation and management of of trust/reputation in open networks [1], [2], [3], [4], [5], [6]. All these methods use semantic information and focus on a static snapshot of a social network, which does not capture all of the communication behavior. Conversely, we study behavioral trust purely from the observed communication statistics, using no semantic information. We give measures of behavioral trust which apply to very rapidly dynamic communication networks, for example the Twitter network. This work is an abridged version of [7].

We adopt the notion of interpersonal trust as in [8], where trust as a social tie between a trustor and a trustee [9].

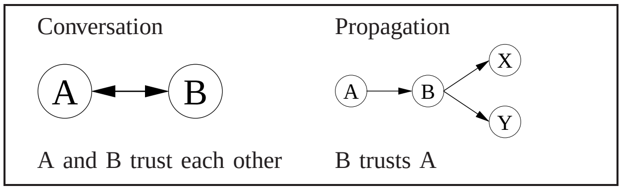
II. Behavioral Trust

Let us formally define the problem. The input is the communications of a social network, a set of *communication 3-tuples*,

$$\langle \mathsf{sender}, \mathsf{receiver}, \mathsf{time} \rangle;$$

We do not use communication content. The output is a behavioral trust graph T induced from these inputs. The nodes in this graph are the senders and receivers. The edge weight w_{ij} is the strength of the trust relationship from node i to node j (trust can generally be an asymmetric relationship).

We propose that trust between two nodes A and B will result in certain typical behaviors. These behaviors are not only an expression of trust, but could also facilitate the development of further trust. The simplest such behavior is conversation. Two people who trust each other are likely to converse; and, continued conversation can enhance their trust relationship. Such behavioral expressions are not guaranteed expressions of trust; they are more noisy indicators of trust. The more often they occur, the more indicative they are of trust. We will focus on two particular behaviors as an expression of trust: conversation and propagation.



A. Conversational Trust

Let A and B be a pair of users, and let $\mathcal{M} = \{t_1, t_2, \ldots, t_k\}$ be a sorted list of times when a message was exchanged between A and B. It is possible to break this list into *conversations* $\mathcal{C} = \{C_1, \ldots, C_\ell\}$, where two messages in the same conversation occur close to each other. We only used conversations of size at least 2 in our experiments, in which case $\mathcal{C}$ may not be a complete partition of $\mathcal{M}$.

The measure of conversational trust will be based on the conversations in $\mathcal{C}$, obeying the following properties:

- Longer conversations implies more trust.
- More conversations implies more trust.
- Balanced participation by A and B implies more trust.

Note that one could add other requirements, for example, if people who did trust each other stop keeping in touch, their trust will likely deteriorate over time - i.e. more spaced apart conversations implies less trust. However, the above three properties are a good starting point.

We define the conversational trust $T_c(A, B)$ as follows:

$$T_c(A, B) = \sum_{i=1}^{l} \|C_i\| \cdot H(C_i)$$

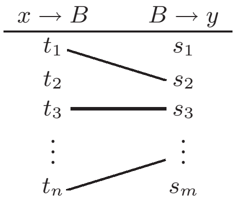
Where $H(C_i)$ is a measure of the balance in the conversation. We use the entropy function to measure balance:

$$H(C_i) = -p \log p - (1 - p) \log(1 - p),$$

where $p(C_i)$ is the fraction of messages in the conversation C_i that were sent by A. One can verify that many, long and balanced conversations lead to high trust as measured by T_c. The complexity of the algorithms for computing conversational trust is $O(|D| \log |D|)$, where $|D|$ is the size of the communication stream.

III. Propagation Trust

Our second measure of trust is based on the propagation of information. If a person A sends a message to person B and if B within some time interval δ propagates the message to some third person X, then we say that B propagated the information received from A. If B propagates information from A often, then we propose that B must be trusting A. As with conversational trust, propagation trust is measured using only statistical communication data without semantic information. Each time B propagates from A, it may be to a different person. Note that this measure of trust (unlike the conversational trust measure) is directed. It is possible for B to be propagating information from A but not vice versa.



To identify potential propagations by B, we need to match messages incoming to B with outgoing messages that are the potential propagations. We use a linear time maximum matching algorithm satisfying a causality constraint developed in [10]. The subset of messages in this maximum matching which were from A are the messages which B propagated from A. We only consider as a valid propagation the pairs (A, B) for which there were a statistically significant number of propagations, as compared to a random communication data stream with the same in and out-degree distributions, as in [10].

Given the valid propagations (A, B), define the quantities: m_{AB}, the number of messages A sent to B; prop_B, the number of propagations by B (the size of the matching above); prop_{AB}, the number of messages A sent to B that were propagated (the subset of the matching containing messages from A. We consider two intuitive ways to measure the directed trust weight $T_p(B, A)$ from B to A:

$$(i)\ T_p(B, A) = \frac{\text{prop}_{AB}}{\text{prop}_B}; \qquad (ii)\ T_p(B, A) = \frac{\text{prop}_{AB}}{m_{AB}}.$$

The first measure captures how much of B's propagation energy is spent propagating messages from A; the second captures the fraction of A's messages B considers worthy of propagating. We have tried both in our experiments, and they yield similar results. We only report the results of (i). In extremely heterogeneous networks, these two measures could capture different aspects of trust, however in homogeneous networks they behave similarly.

IV. Experiments On Twitter Data

Twitter is a popular online free service that enables you to broadcast short messages to your friends or ”followers”, or engage in directed conversations with specific individuals. We constructed a dataset by collecting the publicly available communications between tweeters. The dataset consists of more then 2 million distinct users, of which about 1,910,000 are senders (not all of the users are active). There are about 230,000 public directed messages (“tweets”) per day.

Twitter allows the ability to conveniently and explicitly identify that you are propagating a message through the notion of a *retweet*. Short of interviewing people on who they trust, a retweet (a true propagation) is the next best construct within Twitter for users to explicitly indicate trust in another user. Thus, retweeting gives us a way to validate our behavioral trust measures.

A. Computing Conversation and Propagation Trust Graphs

We used messages over a 10 week period, containing 15,563,120 directed messages and 34,178,314 broadcast messages. We summarize some of the properties of the computed trust graphs, and how they relate to each other.

Node set overlap

	T_c	T_p
T_c	82,947	69,203 (83%)
T_p	69,203(70%)	99,534

Edge set overlap

	T_c	T_p
T_c	202,058	173,638 (86%)
T_p	173,638(70%)	323,820

There is above random similarity between T_c and T_p, indicating that the two trust graphs are capturing similar relationships.

B. *Trust Based Communities in T_c and T_p*

Trust is the foundation of communities, hence we may discover communities as clusters with high within-cluster trust (we use the algorithm in [11]).

	# of Groups	*Max. Group Size*	*Avg. Group Size*
T_c	82947	280	7.06
T_p	81340	316	8.17

The two trust-graphs give similar results, having roughly the same number of communities, as well as a very similar average community size. Indeed this similarity can be more quantitatively measured by comparing the sets of clusters arising from T_c versus T_p. To do this we use the best match method in [12].

	T_c	T_p	Random
T_c	1.00	0.79	0.42
T_p	0.79	1.00	0.43
Random	0.42	0.43	1.00

The trust-based communities from T_c and T_p are more similar than random sets of this same size distribution.

C. *Validating T_c and T_p Using Retweets*

A *retweet* is a definite propagation, indicating trust. Thus, we take $A \longrightarrow B \xrightarrow{\text{retweet}} x$ as a proxy for directed trust edge $B \rightarrow A$ in T_r. For our 10 weeks of Twitter data, T_r had 90,057 nodes and 103,279 directed edges. About 20% of the node set in T_r overlapped with the node sets of T_c and T_p

Our main experimental result is that the behavioral trust graphs do indeed represent trust (as captured by retweets).

T_c vs. T_r

	% edges in T_r
T_c	11.6 %
T_{random}	2.5 %
T_{degree}	2.7 %

T_p vs. T_r

	% edges in T_r
T_p	14.4 %
T_{random}	3 %
T_{degree}	2.9 %

The behavioral trust graphs capture more than 4 times as many retweets as a random null hypothesis (random nodes and their neighborhoods) as well as the prominence (high degree) null hypothesis (high degree nodes and their neighbors). In fact the prominence based trust graph is worse than random!

V. Conclusions

Our main contributions are *measurable* behavioral metrics for dyadic trust. Our results indicate that our behavioral trust measures correlate well with retweets (significantly better than a random null hypothesis), and better than a simple measure of trust based on prominence. The surprising result is that prominence based trust does not fare better than random. These results are preliminary in the sense that there is a lot more information in the behavioral trust graphs than is presented here, and so there are many directions for future work.

1. How does T_c (at higher thresholdeds) relate to T_p.
2. $T_c \cap T_p$ would be also interesting to study, as it provides a more stringent measure of trust.
3. The advantage of statistical algorithms are that they are efficient, but they ignore much information. We may be able to improve the measures with simple semantic analysis.
4. Trust is a contextual relationship – a node may trust one set of nodes in one context (eg. medical advice) and another in another context (eg. movie advice). Semantic analysis of the statistical behavioral trust graphs could be used for context.
5. Efficient algorithms for statistically analyzing the values of messages along different dimensions can considerably enhance the behavioral trust measures (see for example [13] for methods to estimate value of messages).

Acknowledgments

This material is based upon work partially supported by the U.S. National Science Foundation (NSF) under Grant Nos. IIS-0621303, IIS-0522672, IIS-0324947, CNS-0323324, NSF IIS-0634875 and by the U.S. Office of Naval Research (ONR) Contract N00014-06-1-0466 and by the U.S. Department of Homeland Security (DHS) through the Center for Dynamic Data Analysis for Homeland Security administered through ONR grant number N00014-07-1-0150 to Rutgers University. This research is continuing through participation in the Network Science Collaborative Technology Alliance sponsored by the U.S. Army Research Laboratory under Agreement Number W911NF-09-2-0053.

The content of this paper does not necessarily reflect the position or policy of the U.S. Government, no official endorsement should be inferred or implied.

References

[1] T. Beth, M. Borcherding, and B. Klein, "Valuation of trust in open networks," in *Proceedings of ESORICS*, 1994.

[2] V. Buskens, "Social networks and trust," in *The Netherlands: Kluwer Academic Publishers*, 2002.

[3] A. Abdul-Rahman and S. Hailes, "Supporting trust in virtual communities," in *Proc. 33rd Hawaii Int. Conf. on Sys. Sci.*, 2000.

[4] K. Aberer and Z. Despotovic, "Managing trust in a peer2 -peer information system," in *Proc. CIKM 01*, 2001, pp. 310–317.

[5] E. Gray, J.-M. Seigneur, Y. Chen, and C. Jensen, "Trust propagation in small worlds," in *Proc. 1st Int. Conf. on Trust Management*, 2003.

[6] U. Kuter and J. Golbeck, "Sunny: A new algorithm for trust inference in social networks using probabilistic confidence models," in *AAAI*, 2007, pp. 1377–1382.

[7] S. Adali, R. Escriva, M. Goldberg, M. H. M. Magdon-Ismail, B. Szymanski, W. Wallace, and G. Williams, "Measuring behavioral trust in social networks," RPI, Tech. Rep. 10-03, 2010.

[8] K. Kelton, K. R. Fleischmann, and W. A. Wallace, "Trust in digital information," *Journal of the American Society for Information Science and Technology*, 2007.

[9] R. C. Mayer, F. Schoorman, and J. Davis, "An integrative model of organizational trust," *Academy of Management Review*, 1995.

[10] J. Baumes, M. Goldberg, M. Hayvanovych, M. Magdon-Ismail, W. Wallace, and M. Zaki, "Finding hidden group structure in a stream of communications," *ISI*, 2006.

[11] J. Baumes, M. Goldberg, and M. Magdon-Ismail, "Efficient identification of overlapping communities," *ISI*, 2005.

[12] M. Goldberg, M. Hayvanovych, and M. Magdon-Ismail, "Measuring similarity between sets of overlapping clusters," *submitted*.

[13] Y. Zhou, K. Fleischmann, and W. Wallace, "Automatic text analysis of values in the enron email dataset: Clustering a social network using the value patterns of actors," in *Proc. of the 43rd Hawaii Int. Conf. on Sys. Sci.*, 2009.