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 T. Williams

***Abstract*—Trust is an important yet complex and little understood dyadic relation among actors in a social network. There are many dimensions to trust; trust plays an important role in the formation of communities in social networks, in assessing quality and credibility of information as well as in determining how information flows through the network.**

In this paper, we present algorithmically quantifiable measures of trust which can be determined from the communication behavior of the actors in a social communication network. The basis for our study is a proposition that trust results in characteristic communication behavior patterns which are statistically different from random communication in a network. Detecting the statistically significant realizations of this trust-like behavior allows us to develop a quantitative measure of the *who-trusts-whom* relation in the network.

Since our measure of trust is based on quantifiable behavior, we call it behavioral trust. We develop algorithms to efficiently compute behavioral trust and we 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: whether nodes will believe information they receive, choose to transmit it to some other node. 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. In general, when we are deciding whether or not to trust a person, we are all influenced by a host of factors, such as: (1) our own predisposition to trust, which is linked to our psychology, which itself was influenced by various events over our lifetime; these events can be completely unrelated to the person we are deciding to trust or not trust; (2) our relationship and past experiences with the person and with his or her friends, including rumors and gossip; and (3) our opinions of actions and decisions the person has made in the past.

In order to be able to capture and/or quantify trust, we focus on some specific properties of trust, which are simplified, so that these properties can be captured algorithmically. We aim to quantitatively measure dyadic trust (trust between two entities) based on observed communication behavior in social networks – we call this *behavioral trust*. A useful analogy to keep in mind is the saying "imitation is the best form of flattery" – imitation is a behavior which is indicative of some dyadic relationship; similarly, there are behaviors which are indicative of trust.

A typical social network consists of actors (individuals) and some form of communication between them, which could be phone calls, emails, blog posts, etc. Increasingly, a great deal of social relationships take place predominantly in the form of electronic communications. People meet and form trust relationships, participate in activities without any face-to-face contact. As a result, the interactions between individuals in the social network is a good indicator of their social relationships with these individuals. An aspect of trust is based on the notion of embeddedness [1] which shows that the interactions between individuals form a basis from which a trust relationship may grow. Sometimes these interactions may not require trust. However, they establish a relationship that can be used to build trust. The various characteristics of these relationships, such as persistence of communications and the balance in participation, may signal the existence or formation of a trusting relationship.

The social mechanisms with which people form trusting relationships in online communities is a fairly new topic with a lot of unknowns. In this paper, we study a number of social behaviors that take place in this space: conversations and propagation of information from one person to another. We develop *statistical* measures based on the timing and sequence of communications, not the textual content. We give efficient algorithms for computing our measures, making them scalable to social networks on millions of nodes. We show that these two types of behavior correlate strongly with each other in terms of the individuals involved and the communities formed. We also show that they correlate with actual forwarding behavior indicative of trust. These results give us a new set of behavioral measures that can be used to measure existence, emergence or dissolution of trusting relationships in social networks.

Related Work. There has been work done on trust in computer science as well as in social science. In [2], Beth et al. present a method for valuation of trustworthiness in open networks. In [3], Buskens discusses proposes explanations for the emergence of trust in social networks when actors can label others as untrustworthy, and when actors are informed regularly about trustworthy behavior of others. Abdul-Rahman and Hailes [4] and Aberer and Despotovic [5] study reputation based trust and trust management. Abdul-Rahman and Hailes present a model in which agent's tune their measures of trust based on observed reputations, and Aberer and Despotovic discuss a trust model that is grounded in real-world social trust characteristics, and based on a reputation mechanism, or word-of-mouth. Their proposed model allows agents to decide which other agents' opinions they trust the most, and

allows agents to progressively tune their understanding of another agents subjective recommendations. In [5], Aberer and Despotovic present scalable algorithms that require no central control and allow for estimating trust by computing an agents reputation from its interactions with other agents. In [6], Gray, Seigneur, Chen and Jensen develop trust-based security mechanisms using small world concepts to optimize formation and propagation of trust among entities in a massive, networked infrastructure of diverse units. They summarize that, in a very large mobile ad hoc network, trust, risk, and recommendations can be propagated through relatively short paths connecting entities. In [7], Kuter and Golbeck describe a different approach for estimating trust in various computing systems. They give an explicit probabilistic interpretation for confidence in social networks. They describe SUNNY, a new trust inference algorithm that uses a probabilistic sampling technique to quantify confidence and trust. SUNNY computes an estimate of trust based on only those information sources with high confidence estimates.

All the methods proposed above use semantic information in some way and/or focus on a static snapshot of a social network, which does not capture all of the communication behavior and dynamics. Conversely, we study the problem of behavioral trust purely from the observed communication statistics, using no semantic information. We give measures of behavioral trust which apply to rapidly changing dynamic, streaming communication networks, for example the Twitter network.

We adopt the notion of interpersonal trust as proposed by Kelton et al. in [8], which treats trust as a social tie between a trustor and a trustee [9]. Trust develops as part of an emotional relationship between a pair of people akin to the concepts of emotional and relational trust [10], [11].

II. Behavioral Trust

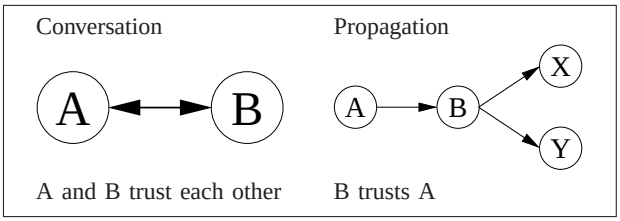
Let us formally define the problem now. The input is the communication stream D of a social network, specified by a set of *communication 3-tuples*,

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

note that we do not use communication content, only the sender-receiver-time data. The output is a behavioral trust graph T induced from these inputs. The participants of the communication are the nodes of this graph. The edges (to be defined below) are weighted; the edge weight w_{ij} indicates the strength of the trust relationship from node i to node j (trust can generally be an asymmetric, directed relationship).

The basis for this work is the observation that trust between two nodes A and B will result in certain typical behaviors. These behaviors are not only an expression of trust, but can also facilitate the development of further trust. The simplest such behavior is just conversation. Two people who trust each other are likely to converse; in addition, continued conversation can lead to an enhancing of their trust relationship. Note that such behavioral expressions are not guaranteed expressions of trust. It is possible to have a conversation with someone who you do not trust; it is also possible to trust someone but not converse with them. Thus, such behavioral expressions of trust should be more viewed as noisy indicators. The more often they occur, the more likely that a trust relationship is likely to exist or to develop. Further, since our measures are statistical, they ignore some of the contextual aspects of trust. For example you trust your doctor for medical advice and your accountant for tax advice. From the behavioral point of view, you would converse with both your doctor and accountant, however, they are distinct forms of trust. The contextual aspect could be added back through the notion of "trust communities" but our present goal is to simply measure whether there is *a* trust relationship between two entities A and B.

Note that it is also possible to measure distrust through typical behavior expressed by distrust. For example, seeking of a second opinion could be considered a measure of distrust. For the scope of this present work, we focus on measuring dyadic trust. We will focus on two particular behavior patterns as an expression of trust: conversation and propagation. Specifically, if two nodes converse, then they are more likely to trust each other; and a prolong conversation reinforces this conclusion. If one node propagates information from another then it suggests that the propagator trusts the information. Similarly, a repeated propagation makes the conclusion stronger.



Our goal is to develop algorithmic measures of conversation and propagation, and validate these as measures of trust in the Twitter network.

A. Conversational Trust

We postulate that the longer and more balanced a conversation is between two nodes, the more likely it is that they have a trust relationship; in addition, the more conversations there are between such a pair of nodes, the more tightly connected they are. The basic task is to first identify when two nodes are conversing.

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 the times when a message was exchanged between A and B. The average time between messages is defined as $\tau = (t_k - t_1)/k$. We would like to split the message set $\mathcal{M}$ into a set of disjoint conversations. To do this, we introduce a user-defined "smoothing" factor S, and say that two consecutive messages t_i, t_{i+1} are in the same conversation if $t_{i+1} - t_i \leq S \cdot \tau$. A straightforward algorithm can be used to construct the set of conversations $\mathcal{C} = \{C_1, \ldots, C_\ell\}$, making a single pass through $\mathcal{M}$ based on the following observation. Suppose we are working on conversation $C = \{t_{i_1}, \ldots, t_{i_c}\}$; if $t_{i_c+1} - t_{i_c} < S \cdot \tau$, then we add t_{i_c+1} to the conversation C,

otherwise we start a new conversation. We only use conversations of length 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 postulates:

- Longer conversations imply more trust.
- More conversations imply 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. Given the stream of communications, we construct the conversation trust graph, $T_c(V, E_c)$, where the weight between a pair of agents $\{A, B\}$ is $T_c(A, B)$; we normalize so that the maximum weight is 1 and only keep edges with weight at least 0.01 (this choice is arbitrary, and leads to roughly the same order of edges as in the propagation trust graph as we describe below). The complexity of the algorithms for computing conversational trust is $O(|D| \log |D|)$, where $|D|$ is the size of the communication stream.

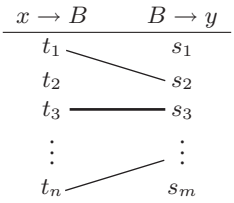
B. Propagation Trust

Our second measure of trust is based on the propagation of information. If A sends a message to B, and if B, within some time interval δ, propagates the message to some third person x, this is indicative of trust. If B propagates information from A often, then we propose that B must be trusting A. Note that finding whether B is propagating information from A is a hard problem even with processing of text – for example, the text from A may be altered as B propagates it. We develop a measure of propagation trust using only statistical communication data, without semantic information, similar to conversational trust. Each time B propagates information from A, it may be to a different person; each such propagation signifies trust in A even though it may be to different people. 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.

We now describe how to get the propagation trust graph $T_p = (V, E_p)$. We need to construct the directed edge $B \to A$, which means that B trusts A. We begin with two sorted time lists: the set of messages incoming to B; and, the set of messages sent by B. We wish to associate pairs of messages (one from the received list, and one from the sent list) as propagations. Based on communication statistics alone, we cannot definitely determine which messages from B are propagations; however, we can identify "potential propagations". Specifically, we say that a message m_1 received by B was potentially propagated by a message m_2 sent by B if their times are close enough to satisfy the propagation constraint:

$$\tau_{\min} \leq t_{m_2} - t_{m_1} \leq \tau_{\max}.$$

To find the maximum number of potential propagations by B, and in particular, the number of A's messages which B potentially propagated, we need to match messages incoming to B with messages outgoing from B. These matches are the potential propagations, as illustrated below.



The first step is to find the maximum number of potential propagations; this corresponds to finding a maximum sized matching, where each match satisfies the propagation constraint. This matching problem can be solved efficiently in linear time [12]. A subset of messages in this maximum matching will be from A; these message pairs are the ones we take as B's (potential) propagations of information from A. We only consider as a valid propagation edges 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 (a similar approach was used in [12]).

Notice that in the matching illustrated above, none of the links cross. This corresponds to a causality constraint, namely that if B propagated two messages which he received at times $t_1 < t_2$, the times of the propagations must also satisfy this ordering. One can show that some maximum matching satisfies this constraint, and infact a greedy matching which starts with the first possible match is one such matching. Given that the maximum matching can be computed in linear time, the entire algorithm to find propagations (which includes sorting the message times) takes $O(|D| \log |D|)$.

Given the valid propagations (A, B), we 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.

Next we discuss the Twitter data followed by experiments to study and validate our trust measures.

III. Twitter Data

Twitter is a popular online free service that enables one to broadcast short messages to ones friends or "followers", or engage in directed conversations with specific individuals. "Tweets" are text-based posts of up to 140 characters displayed on the author's profile page that are delivered to the author's subscribers (followers). Senders can restrict delivery to those in their circle of friends or, by default, allow open access.

We constructed a dataset by collecting the publicly available communications between tweeters. We reduced it into our standard input format (sender, receiver, time). 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 provides a convenient, explicit way to identify that you are propagating a message through the notion of a *retweet*. When we gather retweets, we only gather the information about the original sender of the message and the retweeter. There are two types of retweeting: directed and broadcast: directed retweeting is to a particular receiver, and a broadcasted retweet goes to all followers of the retweeter. Short of interviewing people and asking 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.

IV. Experiments On Twitter Data

We first ran some experiments to compare the conversation and propagation trust graphs. In many aspects, they are similar. We then used Twitter retweets to validate our measures of trust, and we show that our measures fare better than random and prominence based null hypotheses.

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 use only directed messages to identify conversations for the conversation trust graph T_c; for the propagation trust graph T_p, we use directed and broadcast messages (broadcasts are only used for outgoing messages).

To determine the statistically unlikely behavior, specifically, to determine how many propagations are a significant number, we built a random graph model for the Twitter data. The runs with over $M = 1000$ random data sets, showed that four propagations of the form $A \to B \to x$ never happened, which (using standard Chernoff bounds) gives a greater than 99% p-value at the 95% confidence level that four propagations in the Twitter data would not happen under the null hypothesis that Twitter is a random graph without dyadic relationship structure. A summary of some of the properties of the computed trust graphs, and how they relate to each other are in the table below.

T_c	T_p
Smoothing par. $S = 4$	$\tau_{\min} = 1;\ \tau_{\max} = 120$ (min)
202,058 undir. edges	323,820 dir. edges

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

We treat the undirected edges in T_c as two directed edges for purposes of comparing edge sets. We note that there is significant similarity between T_c and T_p, which is significantly above random considering that there are over 2 million users in our data. This says that the type of relationship the two trust graphs are capturing is similar.

B. Trust Based Communities in T_c and T_p

Trust is the foundation of communities, and it should be possible to discover communities in the Twitter network by identifying groups of nodes with a high degree of trust between members of the group. We do it by defining a cluster density in terms of the trust-weights on the edges, and then using local optimality together with the iterative search algorithm for cluser identification as described in [13]. For the sake of simplicity, we treat the graphs as undirected, though the directed clustering method could also be used. Some basic statistics of the communities are shown below.

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

Notice that the two trust-graphs have roughly the same number of communities with 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 [14]. The best match method takes every cluster arising from T_c and compares it with the best match cluster from T_p, and vice versa. The similarity between the two sets of clusters is then the average best match similarity. We can also consider the similarity between the T_c-clusters and a random set of clusters with the same size distribution as the T_p-clusters; this serves as a null distribution for determining whether the observed similarity is significant. We compare the set of trust based communities to 1,000 different random sets of clusters to get an average similarity. The results are shown below.

	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

We see that the trust-based communities coming from T_c and T_p have a similarity larger than would be expected for random sets with the same size distribution. This is a further indication that both the conversational and propagation trust graphs are capturing a similar dyadic relationship.

We have studied some of the properties of the conversation and propagation trust graphs, to establish that though they are measuring different behaviors, both these behaviors result in establishing similar relationships between nodes, both at a local edge and node level, as well as on a collective level as seen through the lens of trust-based communities. Thus, both measures seem to be capturing at least some part of the same phenomenon. We would like to now provide some evidence that this phenomenon is indeed trust.

C. Validating T_c and T_p Using Retweets

A *retweet* is a definite propagation; we make the assumption that when a user propagates information from some other user, there must be some element of trust between the two users. Thus, we take a retweet of the form

$$A \longrightarrow B \overset{\text{retweet}}{\longrightarrow} x$$

as a proxy for directed trust $B \rightarrow A$ (x could be an individual or group of individuals, eg. followers) – thus, we may consider directed as well as broadcasted retweets. A broadcast propagation is not as significant a trust indicator as a directed propagation, since a directed retweet indicates that the user has carefully processed the information and deemed it appropriate to forward to some specific friend. Thus, we consider the broadcast retweets as less significant measures of trust than directed retweets. We therefore build the *retweet-trust* graph T_r as follows. If there is at least one directed retweet $A \rightarrow B \rightarrow x$, then the directed edge $B \rightarrow A$ exists in T_r; if there are at least two broadcast retweets by a node B of two different messages from A, then the directed edge $B \rightarrow A$ exists in T_r. The choice of 1 for the number of directed retweets to indicate trust and 2 for the number of broadcast retweets to indicate trust are somewhat arbitrary and chosen for illustration. 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 (recall that the node sets of T_c and T_p are very similar).

Our main experimental result is that the behavioral trust graphs do indeed represent trust (at least as captured by retweets). Every edge in the behavioral trust graphs T_c and T_p represent a trust relationship. If the retweet graph is our proxy for trust, we should therefore expect that every edge in the behavioral trust graphs should be present in the retweet graph. In fact the fraction of behavioral trust edges which are present in the retweet graph is a measure of how well the behavioral trust is capturing "retweet" trust, which in turn is a proxy for trust. These results are shown in the table below.

Conversational Trust vs. Retweets

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

About 12% of the edges in T_c are also present in the retweet graph. To understand whether this is significant, we consider two alternate null models for building "trust" graphs. The first is just a random model. So we select a set of nodes randomly; the number of nodes we select is exactly the number of nodes in T_c. We now consider all the communications incident with this random set of nodes to construct the random trust graph T_{random}. As can be seen above, only 2.5% of these edges of T_{random} are present in the retweet graph. Another plausible null model for trust is the prominence model. Thus, one might hypothesize that nodes which send many messages (i.e. nodes with high communication degree) might be trusted nodes. Indeed this is the type of hypothesis consistent with preferential attachment type models. So, we construct the high degree graph T_{degree} in a similar way to the random graph. Instead of selecting random nodes, we select the highest degree nodes (the same number as are present in T_c), and the communications incident with these nodes are the edges. As we see above, the high degree nodes are no more trusted (with respect to the edges appearing in the retweet graph) than the random set of nodes. A similar picture arises in the propagation trust graph T_p.

Propagation Trust vs. Retweets

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

We conclude that the fraction of edges in T_c or T_p which appear in the retweet graph is significant when compared to random nodes or the prominent nodes (as measured by communication degree). This means that behavioral trust links are capturing something more sophisticated than simply links to prominent nodes. Several low degree nodes are also picked. This is to be expected as trust is not a phenomenon restricted to voluminous users. The surprising thing is that prominent nodes do not yield better performance than random nodes, and importantly, the behavioral trust measure performs more than 4 times better than random.

V. Conclusions

The main contribution of this paper is to present *measurable* behavioral metrics for trust. In this way we can quantify dyadic trust (a highly complex relationship) through observable communication behavior in social networks. In particular, our behavioral trust measures require only the communication traffic stream (sender, receiver, time), and does not look at semantic content of the messages. We have used Twitter data to illustrate our methods, which can be applied to very dynamic social communication networks. We were able to use retweet data available from Twitter to validate our measures of behavioral trust because retweets are explicit propagations

of information which indicate a trust in the information. 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.

We emphasize that our measures of trust do not use the retweet data (employed for validation purposes only), and so are applicable to general social networks where all one can observe are communications. The advantage of only using statistical communication data (as opposed to semantic data) is that our algorithms are scalable to larger networks (the Twitter data we analyzed contained 2 million nodes). 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. The conversation graph T_c can be thresholded at higher values to yield a much larger graph than the propagation graph T_p. It would be interesting to study the behavior of T_c and its relationship to T_p as we increase this threshold. We believe this relationship is interesting because we hypothesize that conversation is a beginning of a trust relationship and information propagation relies on a pre-existing trust relationship. Thus, we expect conversation trust to precede propagation trust. Hence, it would be very interesting to study how, in the real data, edges in the conversation trust graph T_c transition from low to high weight, and perhaps eventually into propagation trust edges. If this was indeed observed, it would verify the hypothesis.
2. The intersection of the conversation and propagation graphs $T_c \cap T_p$ would be also interesting to study, as it provides a more stringent measure of trust – not only is there conversation but also propagation.
3. The advantage of statistical algorithms are that they are efficient, but they ignore much information. For example after building the statistical propagation trust graph, we have a set of candidate edges. We may now filter these edges using semantic analysis of content to see which edges correspond to real propagations of *information*, as measured by , for example, partial matching of the content. Thus, we would be identifying the "retweets" through semantic information – this is important for networks where the retweet functionality is not available.
4. Trust is a contextual relationship. In our trust graphs, all the trust relationships are homogeneous. In reality, a node may trust one set of nodes in one context (eg. medical advice) and another set in another context (eg. movie advice). Semantic analysis of the statistical behavioral trust graphs could add the context to behavioral trust.
5. Efficient algorithms for statistically analyzing the value and context of a message can considerably enhance the behavioral trust measures (see for example [15] for methods to estimate value of messages). Specifically, if a conversation contains high value content, it is probably a better indicator of trust. Similarly, if a propagation is a propagation of high value information, it is probably an indication of a stronger trust relationship. Thus, value analysis of messages could considerably enhance the behavioral trust measures.

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] S. Zukin and P. DiMaggio, *Structures of Capital: The Social Organization of the Economy*. NY: Cambridge Univ. Press, 1990.
[2] T. Beth, M. Borcherding, and B. Klein, "Valuation of trust in open networks," in *Proceedings of ESORICS*, 1994.
[3] V. Buskens, "Social networks and trust," in *The Netherlands: Kluwer Academic Publishers*, 2002.
[4] A. Abdul-Rahman and S. Hailes, "Supporting trust in virtual communities," in *Proceedings of the 33rd Hawaii International Conference on System Sciences*, 2000.
[5] K. Aberer and Z. Despotovic, "Managing trust in a peer-2-peer information system," in *Proc. 10th Int. Conf. on Information and Knowledge Management(CIKM01)*, 2001, pp. 310–317.
[6] E. Gray, J.-M. Seigneur, Y. Chen, and C. Jensen, "Trust propagation in small worlds," in *Proceedings of the First International Conference on Trust Management*, 2003.
[7] 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.
[8] K. Kelton, K. R. Fleischmann, and W. A. Wallace, "Trust in digital information," *J. Amer. Society for Information Science and Technology*, vol. 59, pp. 363–374, 2008.
[9] R. C. Mayer, F. Schoorman, and J. Davis, "An integrative model of organizational trust," *Academy of Management Review*, 1995.
[10] J. Lewis and A. Weigert, "Trust as a social reality," *Social Forces*, 1985.
[11] D. Rousseau, S. Sitnik, R. Burt, and C. Camerer, "Not so different after all: A cross-discipline view of trust," *Academy of Management Review*, 1998.
[12] J. Baumes, M. Goldberg, M. Hayvanovych, M. Magdon-Ismail, W. Wallace, and M. Zaki, "Finding hidden group structure in a stream of communications," *Intel. and Sec. Inform. (ISI)*, 2006.
[13] J. Baumes, M. Goldberg, and M. Magdon-Ismail, "Efficient identification of overlapping communities," *IEEE Int. Conf. on Intel. and Sec. Inform. (ISI)*, pp. 27–36, May, 19-20 2005.
[14] M. Goldberg, M. Hayvanovych, and M. Magdon-Ismail, "Measuring similarity between sets of overlapping clusters," *submitted to AAAI 2010*.
[15] 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. 43rd Hawaii Int. Conf. on System Sciences*, Kauai, HI.