

Propagation of Trust and Distrust

R. Guha
rguha@us.ibm.com
IBM Almaden Research Center
650 Harry Road
San Jose, CA 95120

Prabhakar Raghavan
pragh@verity.com
Verity Inc.
892 Ross Drive
Sunnyvale, CA 94089

Ravi Kumar
ravi@almaden.ibm.com
IBM Almaden Research Center
650 Harry Road
San Jose, CA 95120

Andrew Tomkins
tomkins@almaden.ibm.com
IBM Almaden Research Center
650 Harry Road
San Jose, CA 95120

ABSTRACT

A network of people connected by directed ratings or trust scores, and a model for propagating those trust scores, is a fundamental building block in many of today's most successful e-commerce and recommendation systems. In eBay, such a model of trust has significant influence on the price an item may command. In Epinions (epinions.com), conclusions drawn from the *web of trust* are linked to many behaviors of the system, including decisions on items to which each user is exposed. We develop a framework of trust propagation schemes, each of which may be appropriate in certain circumstances, and evaluate the schemes on a large trust network consisting of 800K trust scores expressed among 130K people. We show that a small number of expressed trusts/distrust per individual allows us to predict reliably trust between any two people in the system with high accuracy: a quadratic increase in actionable information. Our work appears to be the first to incorporate distrust in a computational trust propagation setting.

Categories and Subject Descriptors

H.3.3 [Information Storage and Retrieval]: Information Search and Retrieval; H.3.5 [Information Storage and Retrieval]: Online Information Services—*Web based services*; G.1 [Numerical Analysis]: Numerical Linear Algebra; G.2 [Discrete Mathematics]: Graph Theory—*Graph algorithms*

General Terms

Algorithms, Experimentation, Measurements

Keywords

Trust propagation, web of trust, distrust

1. INTRODUCTION

The web increasingly impacts the processes used by individuals to express as well as discern preferences among items. A user may turn to the web for information on purchases such as digital cameras, songs, or movie tickets; or for information on much higher impact acquisitions such as houses, jobs, or even mates. As these decisions and the underlying financial processes themselves migrate to the web, there is growing economic motivation to spread information—and sometimes disinformation—through the web. Open standards and a low barrier to publication demand novel mechanisms for validating information. Thus, we see unscrupulous exploitations of the holes in the social fabric of the web: successful manipulation of stocks by teenagers posting on investment boards under assumed personas; posts by product marketers pretending to be customers extolling the virtues of their product; online relationships that turn sour when one partner uncovers dramatic misinformation with respect to age or gender; link spamming of search engines to simulate popularity; and so forth.

One commonly proposed solution to this problem is to build and maintain a *web of trust* either in microcosm (as for an individual web site) or in macrocosm (across the whole web) that would allow users to express trust of other users, and in return would apply the entire web of relationships and trusts to help a user assess the likely quality of information before acting on it. Through such a web of trust, a user can develop an opinion of another user without prior interaction. The goal of this paper is to propose and analyze algorithms for implementing such a web of trust.

Such a network is a fundamental building block in many of today's most successful e-commerce and recommendation systems. On eBay, for instance, a model of trust has significant influence on the price an item may command. While on Epinions, conclusions drawn from the web of trust are linked to many behaviors of the system, including decisions on items to which each user is exposed.

1.1 Approaches to trust propagation

A natural approach to estimate the quality of a piece of information is to aggregate the opinions of many users. But

this approach suffers from the same concerns around dis-information as the web at large: it is easy for a user or coalition of users to adopt many personas and together express a large number of biased opinions. Instead, we wish to ground our conclusions in trust relationships that have been built and maintained over time, much as individuals do in the real world. A user is much more likely to believe statements from a trusted acquaintance than from a stranger. And recursively, since a trusted acquaintance will also trust the beliefs of her friends, trusts may propagate (with appropriate discounting) through the relationship network.

An approach centered on relationships of trust provides two primary benefits. First, a user wishing to assess a large number of reviews, judgments, or other pieces of information on the web will benefit from the ability of a web of trust to present a view of the data tailored to the individual user, and mediated through the sources trusted by the user. And second, users who are globally well-trusted may command greater influence and higher prices for goods and services. Such a system encourages individuals [4] to act in a trust-worthy manner, placing positive pressure on the evolving social constructs of the web. Indeed, social network theory and economics have considered a variety of facets of this general subject [1, 2, 3, 6, 25].

1.2 Introducing distrust

Recent work [14, 21] give a mathematical approach to the propagation of trust, but does not extend to the case in which users may also express distrust. However, experience with real-world implemented trust systems such as Epinions and EBay suggest that distrust is at least as important as trust. In the absence of treatment of distrust in prior work, it is unclear whether a trust score of 0 translates to distrust or to ‘no opinion’; merely shifting all trust scores so that no negative values remain will not address this fundamental issue. Modeling distrust as negative trust raises a number of challenges—both algorithmic and philosophical. For instance, the principal eigenvector of the trust matrix need no longer be real. Another challenge: what does it mean to combine distrusts through successive people in a chain. Perhaps issues like this have been barriers to modeling distrust in trust propagation. One of the main contributions of our paper is to rectify this situation. We devote significant effort to developing an understanding of appropriate models for the propagation of distrust (Section 3.3.1 and Section 3.4). One our findings is that even a small amount of information about distrust can provide tangibly better judgments about how much user i should trust user j (than information about trust alone). However, using distrust information requires care: the unscrupulous may hold hostage the reputation of a reputable citizen of the web.

1.3 Summary of results

Typical *webs of trust* tend to be relatively “sparse”: virtually every user has expressed trust values for only a handful other users. A fundamental problem is using such webs is that of determining trust values for the majority of user pairs for whom we have not explicitly received a trust rating.

Mechanisms for addressing this problem have been studied in economics, computer science and marketing, albeit typically without a computational component. We present a broad taxonomy of schemes for propagation of trust through a network of relationships, and evaluate 81 such schemes

against a large collections of expressed trusts provided by Epinions. To our knowledge, this is the first empirical study on a large, real, deployed web of trust.

We rank different propagation mechanisms mostly from the perspective of predictive accuracy, in the following sense: at a high level, our experiments involve masking a portion of the known trust ratings and predicting these from the remainder. A large website will naturally have to make trade-offs between accuracy and response time. The hope is that a better understanding of what is correct will lead to better approximations to accuracy.

The remainder of the paper proceeds as follows. Section 2 covers related work. Section 3 then describes our algorithms, and the taxonomy of mechanisms that ties them together. Section 4 covers the web of trust we analyze. In Section 5 we provide experimental results comparing the algorithms and draw conclusions about the effectiveness of trust propagation on real-world data.

2. RELATED WORK

A number of different disciplines have looked at various issues related to trust, including the incremental value assigned by people to transacting with a trusted party and how trust affects people’s beliefs and decision making.

Tversky and Kahneman [13] were the amongst the first to study these phenomena in the context of decision making. There is also a substantial body of work on understanding trust in the field of political science ([9, 18, 23]). We draw a number of useful lessons from these fields, especially in assigning semantics to trust statements, but unfortunately, that work is not computational in nature.

There has been considerable work concerning trust in computer science, most of it focused in the area of security. Formal logical models [8, 10] have been used to in the context of cryptography and authentication. PGP ([24]) was one of first popular systems to explicitly use the term “Web of Trust”, though it was not in the context of search or information flows. We believe that the *same kind* of trust relations between agents can be used for a much wider range of applications than just for belief in statements about identity. Gladwell’s popular book “The Tipping Point” [11] studies the way information flows are mediated by the networks of people and their associated trust relations.

There has been substantial work in the business management community on the value of trust. Akerlof’s classic [1] showed the importance of information regarding the quality of a product (or service). Akerlof showed how information, i.e., knowledge about the trustworthiness of a seller, is vital for the functioning of a market. Trust is an important aspect of on-line communities. Armstrong and Hagel [2] posit the importance of trust and community for on-line commerce.

Recently, due to the emergence of e-commerce, there has been work in the area of developing computational models of trust. Ba, et. al. [5] provide a game theoretic approach of trust and conclude that in the presence of an authenticating third party, most utilitarian course of action for a (market) user is to behave honestly. There have been a number of proposed models and empirical studies of the EBay trust model [12, 16, 17, 20, 22, 19]. However, that line of work has not considered models of propagating trust.

In the last few years, a number of researchers have started looking at the problem of propagating trust through networks. Yu and Singh ([25]) propose a framework which, in

contrast to our work, assumes symmetry and arbitrary transitivity. Kamvar, et. al. [14] consider trust propagation in a peer-to-peer environment and provide an approach that is close to ours, without the incorporation of distrust. In general, most of the work on trust propagation has been inhibited by the lack of empirical data. Very recently, Richardson, et. al. [21] develop a “path-algebra” model of trust propagation which is the closest to ours; moreover, like us, they use data from Epinions to validate their algorithms. To our knowledge, these are the only attempts at a comparative analysis of different propagation algorithms based on a real, large, data set. Moreover, none of the above algorithms handle or even attempt to model distrust in any manner.

3. ALGORITHMS

In this section we describe our framework for trust prediction and develop algorithms in this framework.

3.1 The framework

First, we assume a universe of n users, each of which may optionally express some level of trust and distrust for any other user. These values can be viewed as a real-valued matrix; however to keep our development clean we will in fact partition its entries into two matrices, one for trust and the other for distrust. We take T to be the *matrix of trusts*; t_{ij} is the trust that user i holds for user j . The values t_{ij} are assumed to lie between 0 and 1. Similarly, we take D to be the *matrix of distrusts*, in which d_{ij} again lies between 0 and 1. This formulation allows a user to express both trust and distrust with respect to another user.¹ The main goal of our work is to predict an unknown trust/distrust value between any two users, using the entries available in the trust and distrust matrices.

In the following, we will use M generically to represent a matrix of beliefs, either trust, distrust, or a combination. Since our trust propagation steps will (algebraically) be derived from such belief matrices, we will represent propagation steps in terms of M as well.

3.2 Atomic propagation

We now consider a “basis set” of techniques by which the system may infer that one user should trust or distrust another. Consider a user $i \in [n]$. If we have concluded that i trusts j through some means, an atomic propagation will allow us to carry that conclusion one step further, concluding that i trusts someone related to j .

Each element of the basis set extends a conclusion (such as the conclusion that i trusts j) by a constant-length sequence of forward and backward steps in the graph of expressed trusts. We require that any inference regarding trust should be expressible as a combination of elements of this basis set.²

For example, if we have concluded that i trusts j , and an entry in M indicates that j trusts k , then an atomic propagation would allow us to infer that i trusts k ; we refer to this as *direct propagation*. This propagation is expressible

¹In our experiments, all entries are drawn from $\{0, 1\}$, but our algorithms do not require this.

²Generally, the basis elements may be any family of matrix operations using M . We restrict ourselves to sequences of forward and backward steps following non-zero entries of M since these capture a general and natural set of propagations.

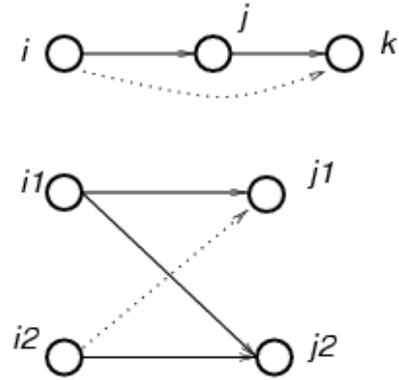


Figure 1: Example of basis elements: Direct propagation and co-citation. The dotted lines indicate trust propagation.

as the matrix M : given any matrix C in which C_{ij} represents current inferences about i 's trust of j , we can replace C with a new inference matrix $C \cdot M$ representing one step of direct propagation. Thus, M is the operator that encodes the direct propagation basis element, as shown in Figure 1.

Another candidate basis element is *co-citation*. For example, suppose i_1 trusts j_1 and j_2 , and i_2 trusts j_2 . Under co-citation, we would conclude that i_2 should also trust j_1 . This basis element is expressed by the matrix $M^T \cdot M$, representing a backward-forward step propagating i_2 's trust of j_2 backward to i_1 , then forward to j_1 (see Figure 1).

The atomic propagations we consider in this paper are described in Figure 2. Let $\alpha = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ be a vector rep-

Direct propagation	M	A trusts B , so trust(A) propagates to B
Co-citation	$M^T M$	A trusts B, C , so trust(B) propagates to C
Transpose trust	M^T	A trusts B , so trust(B) propagates to A
Trust coupling	$M M^T$	A, B trust C , so trust(A) propagates to B

Figure 2: Atomic propagations.

resenting weights for combining our four atomic propagation schemes. Then we can capture all the atomic propagations into a single matrix $R_{M, \alpha}$ as follows:

$$R_{M, \alpha} = \alpha_1 M + \alpha_2 M^T M + \alpha_3 M^T + \alpha_4 M M^T.$$

We now explore how those atomic propagations may be chained together.

3.3 Propagation of trust and distrust

Our end goal is to produce a final matrix F from which we can read off the computed trust or distrust of any two users. In the remainder of this section, we first propose two techniques for computing F from $R_{M, \alpha}$. Next, we complete the specification of how the original trust T and distrust D matrices can be combined to give M . We then describe some details of how the iteration itself is performed to capture two distinct views of how distrust should propagate. Finally, we describe some alternatives regarding how the final results should be interpreted.

3.3.1 Propagation of distrust

As described above, let $R_{M,\alpha}$ be a matrix whose ij th entry describes how trust should be discounted when it flows from i to j via an atomic propagation step; if the entry is 0, then trust does not flow in a an atomic step from i to j . Let k be a positive integer and let $P^{(k)}$ be a matrix whose ij th entry represents the propagation from i to j after k applications of the basis set. In other words, beginning with a belief matrix M , we will arrive at a belief matrix $MP^{(k)}$ after k steps. Thus, the propagation of trust beyond the basis set is expressed as a matrix powering operation.

We give three models to define M (the belief matrix) and $P^{(k)}$ for the propagation of trust and distrust, given initial trust and distrust matrices T and D respectively:

(1) TRUST ONLY: In this case, we ignore distrust completely. The defining matrices then become

$$M = T, \quad P^{(k)} = R_{M,\alpha}^k.$$

(2) ONE-STEP DISTRUST: Assume that when a user distrusts somebody, they also discount all judgments made by that person; thus, distrust propagates only a single step. In this case, we have

$$M = T, \quad P^{(k)} = R_{M,\alpha}^k \cdot (T - D).$$

(3) PROPAGATED DISTRUST: Assume that trust and distrust both propagate together, and that they can be treated as two ends of a continuum. In this case, we take

$$M = T - D, \quad P^{(k)} = R_{M,\alpha}^k.$$

3.3.2 Iterative propagation

We now wish to define F , the final matrix representing the conclusions any user should draw about any other user, based on the computed $P^{(k)}$'s. From our definition, it is clear that each $P^{(k)}$ captures the propagation of trust or distrust via “paths” of length k . In this setting, we present two natural choices.

(1) EIGENVALUE PROPAGATION (EIG): Let K be a suitably chosen (discussed later) integer. Then, in this model, the final matrix F is given by

$$F = P^{(K)}.$$

(2) WEIGHTED LINEAR COMBINATIONS (WLC): Let γ be a constant (that is smaller than the largest eigenvalue of $R_{M,\alpha}$) and let K be a suitably chosen integer. Under this model, F is given by

$$F = \sum_{k=1}^K \gamma^k \cdot P^{(k)}.$$

3.3.3 Rounding

Finally, the result values of F must be interpreted as either trust or distrust. While continuous-valued (rather than discrete-valued) trusts are mathematically clean [21], we work on the assumption that from the standpoint of usability most real-world systems will in fact use discrete values at which one user can rate another. While our mathematical development (like previous work) has been in the continuous domain, we now consider the (non-trivial!) “rounding” problem of converting continuous belief values from

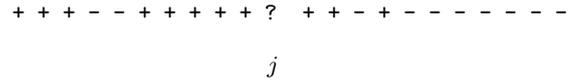


Figure 3: Prediction of j based on the majority of labels of neighbors of i (+ means trust and - means distrust) sorted by the trust scores. Here, the prediction would be +.

an arbitrary range into discrete ones (such as ± 1). This corresponds to applications that demand a Boolean yes/no judgment to the question “Should i trust j ?” This is tantamount to rounding the entries in matrix F to either trust or distrust. We discuss three ways this rounding can be accomplished.

(1) GLOBAL ROUNDING: This rounding tries to align the ratio of trust to distrust values in F to that in the input M . Consider the row vector F_i . We judge that i trusts j if and only if F_{ij} is within the top τ fraction of entries of the vector F_i , under the standard $<$ ordering. The threshold τ is chosen based on the overall relative fractions of trust and distrust in the (sparse) input.

(2) LOCAL ROUNDING: Here, we take into account the trust/distrust behavior of i . As before, we judge that i trusts j if and only if F_{ij} is within the top τ fraction of entries of the vector F_i , under the standard $<$ ordering. The threshold τ is chosen based on the relative fraction of trust vs. distrust judgments made by i .

(3) MAJORITY ROUNDING: The motivation behind this rounding is to capture the local structure of the original trust and distrust matrix. Consider the set J of users on whom i has expressed either trust or distrust. Think of J as a set of labeled examples using which we are to predict the label of a user $j, j \notin J$. We order J along with j according to the entries $F_{ij'}$ where $j' \in J \cup \{j\}$. At the end of this, we have an ordered sequence of trust and distrust labels with the unknown label for j embedded in the sequence at a unique location (see Figure 3). We now predict label of j to be that of the majority of the labels in the smallest local neighborhood surrounding it where the majority is well-defined.

More sophisticated notions of rounding are possible. Notice above that local rounding and majority rounding are “ i -centric”. A j -centric definition is possible in a similar manner. Furthermore, our notion of majority rounding tries to exploit clustering properties. It is possible to derive improved rounding algorithms by using better one-dimensional clustering algorithms.

Our results show that the rounding algorithm is of significant importance in the effectiveness of the system.

3.4 On the transitivity of distrust

It seems quite clear that if i trusts j , and j trusts k , then i should have a somewhat more positive view of k based on this knowledge. In the realm of distrust, however, this transitivity might not hold. Assume i distrusts j , who distrusts k . Perhaps i is expressing the view that j 's entire value model is so misaligned with i 's that anyone j distrusts is more likely to be trusted by i (“the enemy of your enemy is your friend.”) Alternately, however, perhaps i has concluded that j 's judgments are simply inferior to i 's own, and j has concluded the same about k —in this case, i should

strongly distrust k (“don’t respect someone not respected by someone you don’t respect”). We call the former notion *multiplicative* and the latter *additive* distrust propagation.

This problem results because trust and distrust are complex measures representing people’s multi-dimensional utility functions, and we seek here to represent them as a single value. Rather than propose that one answer is more likely to be correct, one can define two corresponding algebraic notions of distrust propagation that may be appropriate for different applications. Notice that by virtue of matrix multiplication, all our earlier definitions implement the multiplicative notion, if we use the trust and distrust values *per se*.

One way to implement the additive distrust notion in our framework is by transforming the matrix M to M' before applying the iteration, as follows:

$$m'_{ij} = \begin{cases} \exp(m_{ij}) & m_{ij} \neq 0, \\ 0 & \text{otherwise.} \end{cases}$$

4. EXPERIMENTAL DATA

We begin with a discussion of Epinions, the provider of our data, and we cover the problems that motivated them to develop and maintain a web of trust between individuals. We then dig into the structure of the graph itself.

4.1 Epinions

Epinions (epinions.com) is a website where users can write reviews about a variety of topics, ranging from consumer durables (such as cars and toasters) to media objects (such as music and movies) to colleges to vacation spots. Given the large number of users (on the order of millions) and the high rate of new reviews (on the order of thousands a day), it is very important to have an automated mechanism for selecting the best reviews for any given topic. A complicating factor in many areas such as movies, music and wines, where tastes are subjective, is that what counts as a good review for one user might not be a useful review for another person.

The reviewable objects are arranged in a taxonomy with top level nodes corresponding to categories of objects (Electronics, Autos, Books, ...). Any user may contribute a review on any object. In addition to a human readable piece of text, each review also typically contains two to five rankings, on various axes (e.g., usability, reliability, etc.) of the object, typically on a scale of one to five. These axes are a function of the kind of object. So, reliability may be an axis for cameras but not for universities. Finally, the user also has to provide an overall rank on a scale of 1–5 for the object.

In addition to writing reviews, a user can also rate reviews of other users on a scale of four ratings, ranging from very useful to useless. Finally, a user can also indicate that s/he ‘trusts’ or ‘distrusts’ another user. Amazon, Slashdot and some other websites also have similar concepts, though they use different terminologies.

Most objects accumulate more reviews than any user can read. Moreover, there is a wide variation in the quality of reviews. Most users are only looking for the top three to five reviews for any particular product. So, given a user and an object, the system needs to identify the top N reviews for that object, for that user.

Often, the user is not researching a particular product

(such as Fizko toaster model 4234) but is instead looking at the page corresponding to the product category (such as toasters or merlots under \$10) and would like some recommendations on which products in that category he should look at. So, given a set of objects (each of which has a number of reviews) and a user, the system needs to identify the top N (typically 5) products to recommend to that user. A variation of this problem is one where we have to pick the top few products to warn the user about (i.e., identify the “lemons”).

Getting one’s reviews rated highly by a number of other users, especially if these users are highly trusted, results in these reviews getting more prominent positions. One complicating aspect at Epinions is that reviewers are paid royalties based on how many times their reviews were read. This motivated many efforts to rig the system, i.e., introduce ratings and trust statements which did not reflect on either the content or the trustworthiness of the user. Distrust was introduced into the system about 6 months after the initial launch, in part to deal with this problem.

Judging by the popularity of the site and the high quality of reviews that are selected, the web of trust seems to be an important and successful mechanism, at least in the context of Epinions.

4.2 Trust graph characteristics

The epinions web of trust may be viewed as a directed graph; the data we obtained consists of 131829 nodes and 841372 edges, each labeled either trust or distrust. Of these labeled edges, 85.29% are labeled trust; we interpret trust to be the real value +1.0 and distrust to be -1.0.

We compute the indegree and outdegree distributions of this directed graph, treating both the trust and distrust edges alike (Figure 4). As in the case of many other statistics on the web, these distributions suggest a power law of exponent -1.7 . Interestingly, this is quite different from that of various power laws that have been observed on the web, where the exponent is generally below -2.0 .

The graph also possesses a large strongly connected component (SCC) with 41441 nodes; the second largest SCC has just 15 nodes. The number of nodes not in the SCC but pointing to it is 39888 and the number of nodes not in the SCC, but pointed to by it is 30823. In other words, the trust graph has a roughly symmetric bow tie structure [7], which shows that the trust graph is well connected even if we use the direction of the edges. If we were to treat the edges as undirected, then we have a giant (weak) connected component with 119130 nodes. We also note that the distributions and overall connectivity properties of the graph are largely preserved even if we restrict our attention to the subgraph induced by the trust edges only.

5. EXPERIMENTS

We now describe our experiments and their results. Based on the algorithmic framework developed in Section 3, our algorithms have the following parameters:

1. Propagation of Distrust (3 cases): Trust only, One-step Distrust, or Propagated Distrust.
2. Iteration Method (3 cases): EIG iteration, WLC iteration, $\gamma = 0.5$, and WLC iteration, $\gamma = 0.9$.
3. Rounding (3 cases): Global, Local, or Majority.

4. Atomic Propagations (3 cases): Direct only ($\alpha = e_1$), Co-citation only ($\alpha = e_2$), or Combined ($\alpha = (0.4, 0.4, 0.1, 0.1)$).

These dimensions result in $3^4 = 81$ experimental categories. (Experiments for the additive distrust model will be presented in the final version of the paper.)

We seek to determine whether any particular algorithm can correctly induce the trust or distrust that i holds for j . Our method is the following. Given the trust graph described above, we remove a single edge (i, j) from the graph, and then ask each algorithm within our taxonomy to guess whether i trusts j ³. Note that even through the matrices T and D are sparse, the final matrix F is not. Considering the dimensions of the matrices involved, it is not feasible to do matrix-matrix multiplications to obtain a matrix of trust scores for every pair of nodes. Instead, we perform a Lanczos-style matrix operation in which, at each step, we do only matrix-vector multiplications. At the end of the matrix-vector multiplications, we obtain a vector that contains the trust score of i for all users. Since all our rounding methods use only this vector, we never need the entire matrix.

We perform this trial on 3250 edges for each of 81 experimental categories, resulting in 263K total trust computations, and tabulate the results in Table 1. In this table, ϵ denotes the prediction error of an algorithm and a given rounding method: the fraction of incorrect predictions made by the algorithm.

As noted earlier, trust edges in the graph outnumber the distrust edges by a huge margin: 85 versus 15. Hence, a naive algorithm that always predicts “trust” will incur an error of only 15%. We nevertheless first report our results for prediction on randomly masked edges in the graph, as it reflects the underlying problem. However, to ensure that our algorithms are not benefiting unduly from this bias, we also take the largest balanced subset of the 3250 trial edges such that half the edges are trust and the other half are distrust. The size of this subset S is 996. We measure the prediction error with respect to this subset and call it ϵ_S . Note that the naive prediction error on S would be 50%. Table 1 shows both values for each experimental category.

5.1 Results

From Table 1, we see that we achieve prediction errors as low as 6.4% on the entire set of 3250 trials and error as low as 14.7% on the subset S . This performance is achieved for the one-step distrust propagation scheme with EIG iteration and $\alpha = (0.4, 0.4, 0.1, 0.1)$.

5.1.1 Basis elements

It was our expectation in undertaking these experiments that direct propagation would be method of choice, and that the other basis elements would perhaps in some limited circumstances provide value. However, the value of co-citation has been proven for web pages by the success of the HITS algorithm [15], so we included it and the other basis elements.

³We insist that i make a Boolean decision about j . This is so that we can measure the efficacy of our algorithms against real data and does not reflect an inadequacy of our algorithm. In fact, as we mentioned earlier, our algorithms operate in the continuous domain and rounding to trust or distrust is the (non-trivial) final step.

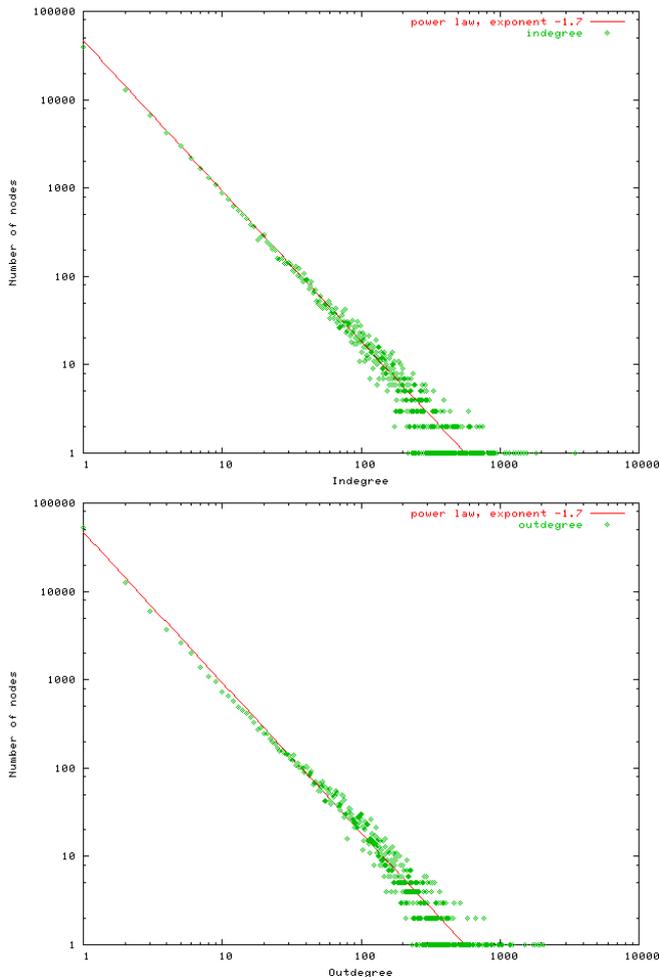


Figure 4: Degree distributions in the trust graph.

Iteration	α	Propagation	Global round.		Local round.		Maj. round.	
			ϵ	ϵ_S	ϵ	ϵ_S	ϵ	ϵ_S
EIG	e_1	Trust only	0.153	0.500	0.123	0.399	0.077	0.175
		One-step distrust	0.119	0.251	0.108	0.223	0.067	0.162
		Prop. distrust	0.365	0.452	0.368	0.430	0.084	0.206
	e_2	Trust only	0.153	0.500	0.114	0.365	0.080	0.190
		One-step distrust	0.097	0.259	0.087	0.234	0.066	0.159
		Prop. distrust	0.149	0.380	0.121	0.279	0.080	0.187
	e^*	Trust only	0.153	0.500	0.107	0.336	0.077	0.180
		One-step distrust	0.096	0.253	0.086	0.220	0.064	0.147
		Prop. distrust	0.110	0.284	0.101	0.238	0.079	0.180
WLC, $\gamma = 0.5$	e_1	Trust only	0.153	0.500	0.123	0.390	0.189	0.163
		One-step distrust	0.093	0.231	0.083	0.205	0.098	0.205
		Prop. distrust	0.102	0.221	0.098	0.199	0.121	0.295
	e_2	Trust only	0.153	0.500	0.113	0.354	0.074	0.174
		One-step distrust	0.088	0.254	0.080	0.231	0.093	0.187
		Prop. distrust	0.126	0.336	0.100	0.252	0.076	0.177
	e^*	Trust only	0.153	0.500	0.108	0.340	0.078	0.159
		One-step distrust	0.086	0.247	0.076	0.217	0.092	0.190
		Prop. distrust	0.087	0.237	0.079	0.203	0.074	0.162
WLC, $\gamma = 0.9$	e_1	Trust only	0.153	0.500	0.123	0.391	0.132	0.152
		One-step distrust	0.102	0.241	0.092	0.216	0.069	0.171
		Prop. distrust	0.111	0.238	0.106	0.211	0.101	0.227
	e_2	Trust only	0.153	0.500	0.113	0.356	0.078	0.184
		One-step distrust	0.092	0.260	0.082	0.235	0.071	0.173
		Prop. distrust	0.134	0.355	0.106	0.261	0.078	0.188
	e^*	Trust only	0.153	0.500	0.107	0.337	0.075	0.169
		One-step distrust	0.091	0.253	0.082	0.222	0.072	0.171
		Prop. distrust	0.091	0.254	0.081	0.209	0.078	0.177

Table 1: Prediction of various algorithms. Here, $e^* = (0.4, 0.4, 0.1, 0.1)$, $K = 20$.

The results, shown in Figure 5, were quite surprising: propagation based only on co-citation alone (basis vector e_2 in the figure) performed quite well. Notice that in this model, simple edge transitivity in the underlying trust graph does not apply: just because i trusts j and j trusts k , we can conclude nothing about i 's view of k . So it is quite surprising that this method performs well. The fact that over all cases in our large table, e^* is the best overall performer seems to indicate that there is a certain amount of resilience to variations in the data by adopting many different mechanisms to infer trust relationships. We recommend this scheme in environments where it is affordable.

5.1.2 Incorporation of distrust

One-step distrust propagation is the best performer with the EIG type of iteration for each of the nine cases (three rounding methods and three basis vectors α). We can consistently recommend one-step distrust in this case. With the WLC type of iteration, distrust is clearly helpful, but depending on the basis vector α , either one-step or propagated distrust may perform better, as shown in Figure 6. The $\gamma = 0.9$ case, which favors long paths, performs worse for one-step distrust than the $\gamma = 0.5$ case. For other distrust models, though, the results are mixed. The most striking result of the figure is that direct propagation (the e_1 case) is the only situation in which distrust actually hurts, sometimes quite substantially; in all other cases we recommend using one-step distrust as robust, effective, and easy to compute. Direct propagation ($\alpha = e_1$) in tree-structured

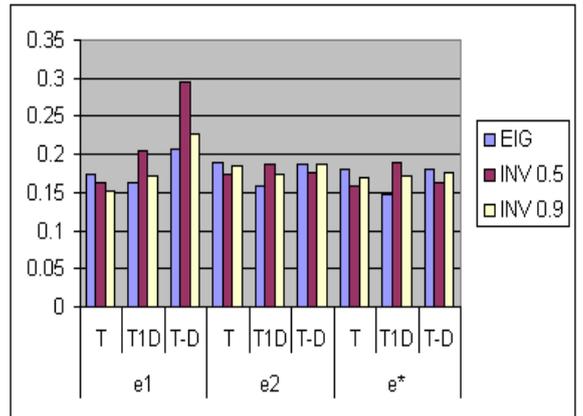


Figure 5: Results for different values of α , majority rounding, against result score ϵ_S .

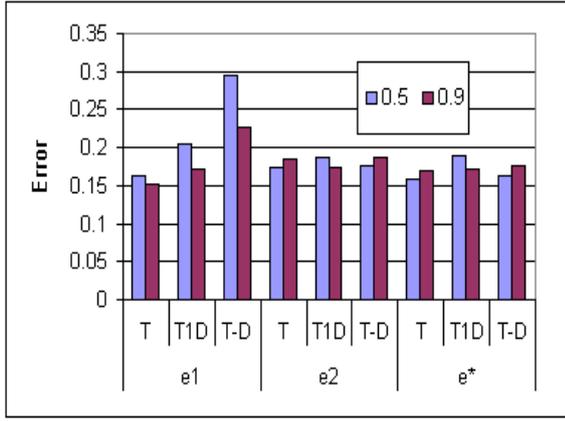


Figure 6: Results for the WLC iteration, $\gamma \in \{0.5, 0.9\}$, showing iteration methods and basis vectors against result score ϵ_S .

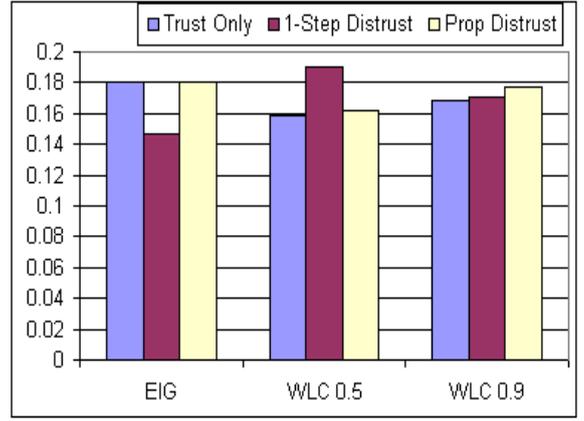


Figure 8: Results for all iteration methods with $\alpha = e^*$, majority rounding, against result score ϵ_S .

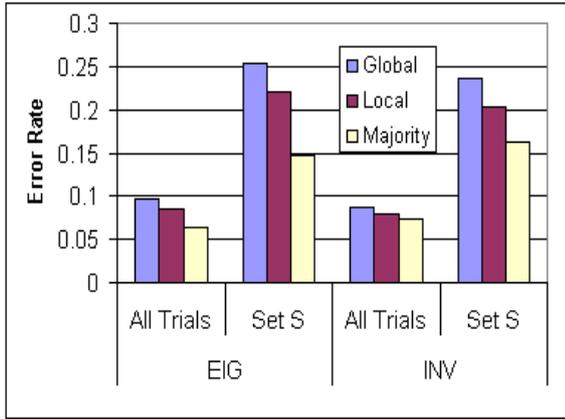


Figure 7: Results for rounding using the best overall settings for the EIG and the WLC iteration against result score ϵ_S .

networks that have no self-loops and no short cycles may result in local information having little impact on the trust scores, which could be undesirable. Recall that the EIG iteration does not introduce any “restart” probability; this would be easy to add, and would result in an algorithm more similar to the WLC iteration.

5.1.3 Rounding

The results for rounding algorithms for the best setting for the EIG iteration (one-step distrust with $\alpha = e^*$) and the best setting for the WLC iteration (propagated distrust, $\gamma = 0.5, \alpha = e^*$). In all cases, majority clustering beats local rounding, which in turn beats global rounding. To our surprise, this part of the algorithm turned out to be quite critical, both in getting good results, and in providing strong performance across all the different cases. We recommend using a decision method like majority rounding.

5.1.4 Iteration models

Figure 8 restricts attention to the generally best basis vector ($\alpha = e^*$) and the best rounding method (majority rounding), and compares results for EIG, and WLC with one-step distrust. Generally, one-step distrust minimizes the impact that a few dangerous trust scores might have, relative to propagated distrust; but as this case shows, it is not always superior. Generally, both methods, and both settings of γ , seem to provide reasonable results; the correct choice may depend on the particular application.

5.1.5 The effect of the number of iterations, K

The following table (Table 2) shows the effect of the number of iterations for three selected settings of parameters. For trust only propagation with $\alpha = e_1$, meaning only direct propagation allowed, increasing the number of iterations has a more dramatic effect on improving the prediction error than for other propagation methods. This is as expected as direct propagation occurs along the directed edges of the graph. In contrast, the other propagation methods, assisted by $\alpha = e^* = (0.4, 0.4, 0.1, 0.1)$, do not enjoy similar dramatic improvements with increasing the number of iterations. In part, this is because the shortest path between most test pairs has length 2, so longer iterations may fail to help.

6. CONCLUSIONS

Over the last few years, a number of ecommerce related sites have made a trust network one of their cornerstones. Propagation of trust is a fundamental problem that needs to be solved in the context of such systems. In this paper, we develop a formal framework of trust propagation schemes, introducing the formal and computational treatment of distrust propagation. We also develop a treatment of “rounding” computed continuous-valued trusts to derive the discrete values more common in applications. Each of our methods may be appropriate in certain circumstances; we evaluate the schemes on a large, real world, working trust network from the Epinions web site. We show that a small number of expressed trusts per individual allows the system to predict trust between any two people in the system with high accuracy. We show how distrust, rounding and other

Iter.	Trust only $\alpha = e_1$		One-step distrust $\alpha = e^*$		Prop. distrust $\alpha = e^*$	
	ϵ	ϵ_S	ϵ	ϵ_S	ϵ	ϵ_S
1	0.120	0.300	0.096	0.209	0.080	0.209
2	0.189	0.216	0.086	0.197	0.082	0.191
3	0.177	0.184	0.088	0.203	0.074	0.184
4	0.157	0.153	0.091	0.206	0.084	0.188
5	0.150	0.156	0.086	0.200	0.082	0.197
6	0.141	0.153	0.086	0.203	0.080	0.197
7	0.135	0.156	0.082	0.197	0.081	0.194

Table 2: Effect of number of iterations on ϵ and ϵ_S for cluster rounding. The iteration type is EIG with $\gamma = 0.9$ and the number of samples is 1000.

such phenomenon have significant effects on how trust is propagated.

7. ACKNOWLEDGMENTS

The authors would like to thank Epinions (<http://www.epinions.com>) for graciously making available the data for this study. We would in particular like to thank Nirav Tolia and Joel Truher for all their help. We would also like to thank Naval Ravikant, Benchmark Capital and August Capital for helping create Epinions.

8. REFERENCES

- [1] G. Akerlof. The market for lemons: Quality uncertainty and the market mechanism. *Quarterly Journal of Economics*, 84:488–500, 1970.
- [2] A. Armstrong and J. Hagel III. The real value of online communities. *Harvard Business Review*, pages 134–141, May-June 1996.
- [3] C. Avery, P. Resnick, and R. Zeckhauser. The market for evaluations. *The American Economic Review*, 89:564–84, June 1999.
- [4] S. Ba and P. Pavlou. Evidence of the effect of trust building technology in electronic markets: price premiums and buyer behavior. *MIS Quarterly*, 26(3):243–268, September, 2002. Study to show that trust can reduce risks for one-time transactions and thus lead to premiums for higher-reputation players.
- [5] S. Ba, A. B. Whinston, and H. Zhang. Building trust in online auction markets through an economic incentive mechanism. *Decision Support Systems*, 2002. Trust through a TTP to aid in building reputation over time. Economic incentives to remain trustworthy. Game theoretic analysis.
- [6] T. Beth, M. Borcharding, and B. Klein. Valuation of trust in open networks. In *3rd European Symposium on Research in Computer Security*, pages 3–19, 1994.
- [7] A. Z. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, and J. Wiener. Graph structure in the web. *WWW9/Computer Networks*, 33(1–6):309–320, 2000.
- [8] M. Burrows, M. Abadi, and R. Needham. A logic of authentication, from proceedings of the royal society, volume 426, number 1871, 1989. In *William Stallings, Practical Cryptography for Data Internetworks*. IEEE Computer Society Press, 1996.
- [9] J. Coleman. *Foundations of Social Theory*. Harvard University Press, Cambridge, Mass., 1990.
- [10] U. Freundrup, H. Httel, and J. Nyholm. Modal logics for cryptographic processes.
- [11] M. Gladwell. *The Tipping Point, How Little Things Can Make a Big Difference*. Little Brown, February 2000.
- [12] D. Houser and J. Wooders. Reputation in auctions: Theory, and evidence from eBay. Technical report, University of Arizona, 2000.
- [13] D. Kahneman, P. Slovic P., and A. Tversky. *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge University Press, April 1982.
- [14] S. D. Kamvar, M. T. Schlosser, and H. Garcia-Molina. The eigentrust algorithm for reputation management in P2P networks. In *WWW*, 2003.
- [15] J. M. Kleinberg. Authoritative sources in a hyperlinked environment. *Journal of the ACM*, 46(5):604–632, 1999.
- [16] P. Kollock. The production of trust in online markets, in advances in group processes. *JAI*, 16:99–123, 1999.
- [17] C. G. McDonald and V. C. Slawson Jr. Reputation in an internet auction model. Technical report, University of Missouri-Columbia, 2000.
- [18] B. Misztal. *Trust in Modern Societies: The search for the Bases of Social Order*. Polity Press, Cambridge MA, 1996.
- [19] P. Resnick and R. Zeckhauser. Trust among strangers in internet transactions: Empirical analysis of eBay’s reputation system. Technical report, University of Michigan, 2001.
- [20] P. Resnick, R. Zeckhauser, E. Friedman, and K. Kuwabara. Reputation systems. *Communications of the ACM*, 43:45–8, 2000.
- [21] M. Richardson, R. Agrawal, and P. Domingos. Trust management for the semantic web. In *Proceedings of the Second International Semantic Web Conference*, 2003.
- [22] J. M. Snyder. Online auction fraud: Are the auction houses doing all they should or could to stop online fraud. *Federal Communications Law Journal*, 52:453–472, 2000.
- [23] P. Sztompka. *Trust. A Sociological Theory*. Cambridge University Press, 1999.
- [24] The MIT PGP Team.
- [25] B. Yu and M. P. Singh. A social mechanism of

reputation management in electronic communities. In *Cooperative Information Agents*, pages 154–165, 2000.