

# Augmenting Reputation-based Trust Metrics with Rumor-like Dissemination of Reputation Information

Sascha Hauke, Martin Pyka, Markus Borschbach, Dominik Heider

► **To cite this version:**

Sascha Hauke, Martin Pyka, Markus Borschbach, Dominik Heider. Augmenting Reputation-based Trust Metrics with Rumor-like Dissemination of Reputation Information. Kai Rannenberg; Vijay Varadharajan; Christian Weber. 25th IFIP TC 11 International Information Security Conference (SEC) / Held as Part of World Computer Congress (WCC), Sep 2010, Brisbane, Australia. Springer, IFIP Advances in Information and Communication Technology, AICT-330, pp.136-147, 2010, Security and Privacy - Silver Linings in the Cloud. <10.1007/978-3-642-15257-3\_13>. <hal-01054521>

**HAL Id: hal-01054521**

**<https://hal.inria.fr/hal-01054521>**

Submitted on 7 Aug 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# Augmenting Reputation-based Trust Metrics with Rumor-like Dissemination of Reputation Information

Sascha Hauke<sup>1,2</sup>, Martin Pyka<sup>3</sup>, Markus Borschbach<sup>1</sup>, and Dominik Heider<sup>4</sup>

<sup>1</sup> Fachhochschule der Wirtschaft in Bergisch Gladbach,  
Hauptstr. 2, 51465 Bergisch Gladbach, Germany  
{sascha.hauke, markus.borschbach}@fhdw.de

<sup>2</sup> Institute of Computer Science, University of Münster,  
Einsteinstr. 62, 48149 Münster, Germany

<sup>3</sup> Department of Psychiatry, University of Marburg,  
Rudolf-Bultmann-Str. 8, 35039 Marburg, Germany  
martin.pyka@med.uni-marburg.de

<sup>4</sup> Center for Medical Biotechnology, University of Duisburg-Essen,  
Universitätsstr. 1-5, 45117 Essen, Germany  
dominik.heider@uni-due.de

**Abstract.** Trust is an important and frequently studied concept in personal interactions and business ventures. As such, it has been examined by multitude of scientists in diverse disciplines of study. Over the past years, proposals have been made to model trust relations computationally, either to assist users or for modeling purposes in multi-agent systems. These models rely implicitly on the social networks established by participating entities (be they autonomous agents or internet users). At the same time, research in complex networks has revealed mechanisms of information diffusion, such as the spread of rumors in a population. By adapting rumor-spreading processes to reputation dissemination in multi-agent systems, this paper shows the benefit of augmenting an existing trust model with pro-actively, socially filtered trust information.

**Key words:** Trust and Reputation, Rumor-spreading, Trust Model

## 1 Introduction

This paper represents part of our work directed at developing a distributed recommendation system of (semi-) autonomous agents, aiding users in determining trustworthy service partners. We envision this system to operate on existing social structures, as, for instance, computationally represented in online social communities. Leveraging reputation-based computational trust and real-world derived social connections as a soft security mechanism, we aim at increasing overall (system) reliability in computer-mediated human interactions within cybercommunities.

In the course of this paper, we will focus on aspects of reputation-based computational trust. In recent literature, two principal views of computational trust

can be distinguished – a cognitive [3, 4] and a probabilistic interpretation. Due to the complexity of accurately representing an entity’s mental state in computational adaptations, computational realizations of cognitive trust are difficult to model.

The probabilistic view of trust, held, among others, by [1, 14, 29], relies on observable data for deriving an, albeit subjective [10, 20], probability with which some entity will perform a particular action in order to establish a trust rating. By employing observed information from the past to predict behavior in the future, trust establishment thus becomes a data driven process. It is well-suited to computational modeling. Recently proposed computational trust models (for reviews see [17, 24]) typically adopt this approach. As our focus is on agent-based models that may operate on social network structures, we consider systems such as ReGreT [23], which features a social dimension, or FIRE [14], itself based on the referral system for multi-agents presented in [29], to be the most related.

Deriving trust is thus based on *reliable* information about the actions expected to be performed by the trusted party. In society, this information is usually procured in two different ways: Either through personal experience derived from prior direct interactions with another entity or via the reputation of an entity as reported by other members of the society (for the time being, we will not consider notions such as role-based or institutional trust). In models, the two first is normally classified as *direct experience*, the latter as *witness information*. To an entity, its direct experiences are the most reliably assessed source of reputation information, yet, it is also the scarcest. Witness information is more abundant, yet its reliability is difficult to assess for an entity. In trust models, witness information is typically communicated in the form of recommendations, involving at least three separate entities: A recommender, a recommendee (whose trustworthiness is to be evaluated) and the evaluating entity, receiving the recommendation.

Assessing the reliability of a recommendation involves evaluating how much trust one can put in the recommender making an accurate recommendation. Trust is a uni-directional, dyadic, non-transitive relation between entities [1]. Thus, an entity’s certainty in the correctness of a recommendation decays rapidly once made by a non-neighboring recommender. Relying only on direct recommendations from trusted neighbors has advantages regarding the reliability of the reputation information. However, this forgoes a potential wealth of additional information. Mui [22] has proposed the establishment of (parallel) recommendation chains between non-neighboring entities. This process, however, suffers from distance effects for long chains and problems when determining the reliability of a particular chain (particularly when determining weighing factors). Bayesian aggregation is also unfit for establishing the reputation of remote agents [22].

## 2 Approach

In recent publications [11, 12], we have proposed a rudimentary mechanism designed to better leverage the information diffusion qualities of social networks

in order to spread reputation information. Aside from increasing the speed of information dissemination [11], the active propagation of reputation information can also serve as a social filter, making it easier to evaluate the validity of the propagated data.

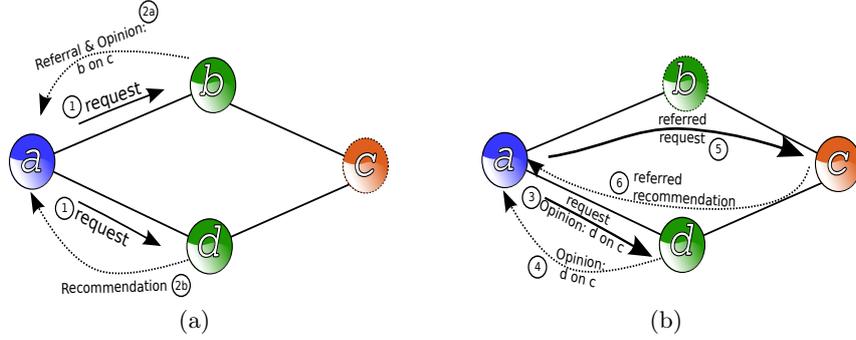
## 2.1 Direct Witness Information

Witness information in typical multi-agent trust frameworks [22, 23, 28] is normally gathered on demand by an entity. When the need for information about a potential interactor arises, a request for reputation information is issued by the requesting entity (*requestor*) to trusted neighbors. If a neighbor holds an opinion on the potential interactor and wishes to share that information, it will return a recommendation to the requestor. Regarding their reliability, recommendations from trusted neighbors can be assessed by evaluating the trust the requestor has in the recommending neighbor. The direct neighborhood the requestor and the recommender affects not only the reliability of a recommendation, as perceived by the requestor, but also implies short achievable response times and low number of exchanged messages. These factors result in relatively *hard* information, suited to by-request communication.

Due to the structure of social acquaintance networks underlying and formed by the trust relations involved in recommending, it is also reasonable to consider one-hop referral recommendations to be hard. One quality of social networks is community structure, expressed through a high average clustering coefficient [16, 26]. The clustering coefficient is a measure of the percentage of *transitive triplets* present in a network [15, 25]. For the sake of illustration, consider three network nodes  $i, j$  and  $k$ .  $i$  is connected to  $j$ ,  $j$  in turn is connected to  $k$ . A transitive triplet is present if  $i$  is also connected to  $k$ . The clustering coefficient is the fraction of transitive triplets to all possible connections that occur throughout the network. The average clustering coefficient [15, 26] is a variation of this measure commonly used in empirical studies.

High average clustering signifies community structure in networks. Within such a community, the likelihood that two entities share common neighbors is higher than in corresponding random graphs [16]. Thus, in a highly clustered network, establishing a 1-hop recommendation chain between an entity  $a$  and a non-neighbor entity  $c$ , facilitated by intermediary entity  $b$ , which is neighboring both  $a$  and  $c$ , has a higher chance of being duplicated via another intermediary entity  $d$ . By establishing multiple parallel recommendation chains of a very short path length, the evaluating entity  $a$  gains the ability to assess the reliability of the remote recommender  $d$  by starting reputation gathering on  $d$ . Due to the distance limitation of 1 hop and high clustering, pertinent information on  $d$  can be gathered from nodes neighboring  $a$ , with the limitation that clustering allows for the establishment of parallel 1-hop chains between  $a$  and  $c$ . While the overhead is considerably higher than 0-hop recommendations (cf. fig. 1), involving a separate trust establishment process for each referral, establishing the reliability of such a recommendation does not suffer from long-distance effects, remains localized close to and is mediated by entities directly known to the requestor.

Furthermore, in case of an inaccurately referred recommendation, both the referring neighbors and the remote recommender can be held directly responsible by reducing their trust value held by the requestor. This also provides an incentive to the referrer to refer only trusted and good-quality recommenders.



**Fig. 1.** Initial message exchange and referred recommendation, on-request opinion provisioning.

We propose to categorize these kinds of witness information as *direct witness information*. It is comparatively easy to assess in terms of reliability and the originating entity of a recommendation is easily identifiable. Furthermore, the amount of communication overhead required for obtaining the information warrants its use in on-demand reputation provisioning. Together with prior direct experiences made by the requesting entity, it forms the basis upon which the *strong reputation* [11, 12] of a potential interactor is computed.

## 2.2 Remote Witness Information

Even when including direct witness information, the amount of information available to an entity attempting to make a trust decision may be scarce. In far-flung social networks, exhaustive breadth-first search for more information is prohibitively expensive time-wise, particularly if the required data cannot be located in the requestor's intermediate neighborhood. Depth-first search leads to quickly deteriorating reliability of gathered information. In order to alleviate the problems with time-criticalness of the reputation information gathering process and rapid decay of reliability, we propose a two-pronged approach.

Firstly, make the information widely available and at the same time decouple the provisioning process from the commonly used on-request mechanism. Secondly, harness the size and structure of the social recommendation network to aid entities in assessing the quality, reliability and relevance of remotely issued reputation information.

### 2.3 Decoupling the Provisioning Process

On-request provisioning of reputation information guarantees that an entity receives the most current information available on a potential interactor. For this mechanism, however, the cost in time and communication overhead are inversely proportionate to the reliability of the information, as determinable by the requestor. An increasing complexity is associated with a decrease in the quality of the acquired data. Figure 1 shows standard witness information gathering and 1-hop recommendation referral.

Therefore, we propose a pro-active propagation of reputation information that mimics the spread of rumors through social networks. Because information that originates beyond the direct neighborhood of an entity is difficult to verify by the entity, it factors less prominently in the trust establishment process in conventional trust models than direct witness information (e.g. [14, 22]); this leads to a diminishing gain in additional information. Consequently, less relevant information, which is thereby less important for trust establishment, may be relegated to a secondary dissemination mechanism that operates on a different time scale than the on-request method.

Fundamentally, the secondary mechanism adapts epidemiological and rumor spreading models operating on complex networks. Various such models have been proposed and investigated, mainly in the fields of sociology [7, 19] and statistical physics [5].

Following established terminology [6], each network element in rumor spreading models is a member of one of three classes, corresponding to ignorant, stifter and spreader nodes. Ignorants are nodes that have not been exposed to a particular rumor and are still susceptible to the information. Spreaders are those that have been exposed to a rumor and are actively propagating the information, while stiflers are privy to a rumor but have ceased spreading it. When a spreader meets an ignorant, the latter turns into a spreader with probability  $\lambda$ . When a spreader meets another spreader or stifter, it turns into a stifter with probability  $\alpha$ . Moreno et al. [21] provide an analysis of rumor mongering in complex networks, presenting time profiles for the propagation process. The dynamics of rumor spreading in networks, as reported, for instance, by [2, 8, 21], suggest sufficient diffusive quality to warrant their application to reputation-based trust metrics.

The basis of our pro-active dissemination mechanism is formed by entities publishing opinions to the network, rather than only providing recommendations to requestors. In order to maintain the added value of those opinions and to avoid overwhelming communication channels, they should be issued only in exceptional circumstances. Such circumstances are constituted, for instance, by an interaction experience significantly deviating from the expected mean, prolonged above (or below) average performance or highly varying, erratic behavior on the part of an interactor.

A published opinion includes an assertion about another entity,  $y$  (a former interactor), time-stamped and signed by the publishing entity  $x$  (the originator). The assertion bundles information regarding the interactor, its trust rating and

perceived reliability of that rating, along with situational data, with situation-dependence denoted by parameter  $\beta$ .  $x$  proceeds by communicating its published opinion on  $y$  to its neighbors. Each publishing entity can only have one published opinion per subject and situation, although the entity can issue a replacement opinion.

$$PubOp_\beta(x, y) = \{Sig_x(Assert_\beta(x, y), timestamp)\} \quad (1)$$

A receiving entity  $r$  confirms reception to the sending entity (either the publishing entity or an intermediate spreader), informing the sender whether or not  $r$  was ignorant to the particular published opinion. If  $r$  does not hold a positive opinion of the sender, it immediately stifles further propagation; otherwise, it evaluates the message, determining whether becoming a spreader or a stifter.

During reception of a particular published opinion, the decision to propagate is primarily dependent on four factors: The prior experience of  $r$  with the sending entity, the difference of the published opinion from the opinion  $r$  has of  $y$ , temporal decay, and the general community agreement regarding  $PubOp_\beta(x, y)$ . According to these parameters, the receiving entity decides, following a rule-based approach, if and with what priority to spread the published opinion. While the first three factors are common values found in trust metrics, we will describe an approach for determining agreement in the following.

#### 2.4 Establishing Reliability through Community Agreement

In order to assign a reliability value to any one published opinion, it is not sufficient for an evaluating entity to just know the reliability of the published opinion's latest propagator, because, generally, that propagator is not the originator of the published opinion. Due to the difficulties in assessing the reliability of a remote originator, as outlined above, we will focus instead on determining the reliability of the content of the published opinion based on a voting process.

Whenever an entity receives and evaluates a published opinion for the first time, it adds a signed token to the assertion included in the message. In this token, the entity includes a vote, denoting whether it agrees, disagrees or has no opinion of its own on the data contained within the assertion.

In order to assess the reliability, we propose the use of Krippendorff's  $\alpha$ -coefficient, a standard reliability measure that can be used regardless of the number of observers, levels of measurement, sample sizes, and presence or absence of missing data [13]. In order to include the total number of informative, i.e. actively agreeing or disagreeing, votes  $N$ , we multiply  $\alpha$  by a monotonously growing scaling function  $f : \mathbb{N} \mapsto [0, 1]$ , accounting for an agent's need to require a certain number of votes on which the reliability measure is based.

$$Rel(Assert_\beta(x, y)) = \begin{cases} f(N) \cdot \alpha & \text{if } \# \text{ agreeing} > \# \text{ disagreeing} \\ -1 \cdot f(N) \cdot \alpha & \text{else} \end{cases} \quad (2)$$

#### 2.5 Transition from Spreader to Stifler State

In classical rumor spreading models [7], transition from spreader to stifler occurs probabilistically once a spreader encounters another spreader or stifler. For our

mechanism, we propose the following procedure: When an entity  $p$  propagates a published opinion, the receiver  $r$  responds by returning its current status as ignorant or non-ignorant. If its status is non-ignorant, i.e. it is a spreader or stifler, it returns the version of the relevant published opinion it is privy to. From this answer, the propagator determines the  $\Delta_{p,r}Rel(Assert_\beta(x,y))$  by first compositing the propagated assertion and the response, followed by calculating the difference in reliability. Compositing is achieved by forming the union of the sets of signed, informative voting tokens contained in the assertions.

$$Rel((Assert_\beta(x,y))^{p \cup r}) = Rel((Assert_\beta(x,y))^p \cup (Assert_\beta(x,y))^r) \quad (3)$$

$$\Delta_{p,r}Rel(Assert_\beta(x,y)) = Rel((Assert_\beta(x,y))^{p \cup r}) - Rel((Assert_\beta(x,y))^p) \quad (4)$$

If  $|\Delta_{p,r}Rel(Assert_\beta(x,y))|$  is smaller than some threshold parameter  $T_{\Delta Rel}$  (meaning that the data provided by  $p$  to  $r$  has little to no effect on the reliability of the published opinion already known to receiver  $r$ )  $p$  will in the future stop propagating  $PubOp_\beta(x,y)$  to  $r$ . Furthermore, it will probabilistically transfer into the stifier state with probability  $p_{stifler}$ .

## 2.6 Evaluating and Integrating Remote Witness Information

For harnessing the social control mechanism in community agreement, the information has still to be integrated into the trust establishment protocol. Most trust models, e.g. [14, 22, 23, 28], already provide compositing mechanisms when dealing with different types of reputation information, such as direct experience, witness reputation or role-based trust. Prototypically choosing the FIRE trust model [14], rumor-like information can be easily integrated into the trust model. FIRE relies on a generic trust formula to calculate a trust value for each of its components:

$$\mathcal{T}_K(a,b,c) = \frac{\sum_{r_i \in \mathcal{R}_K(a,b,c)} \omega_K(r_i) \cdot v_i}{\sum_{r_i \in \mathcal{R}_K(a,b,c)} \omega_K(r_i)} \quad (5)$$

A trust value is thus calculated as the sum of all available ratings weighted by the rating relevance and normalized to  $[-1, 1]$ . Adaptation of remote witness information, spread in a rumor-like manner, is achieved as follows:  $\mathcal{R}_K$  is the set of all published opinions on the entity to be evaluated – here called  $b$  – under a term  $c$ .  $a$  represents the evaluating entity,  $c$  a *term*, under which the evaluation takes place, represented by  $\beta$  in eq. 1. Let  $v_i$  be the rating of a subject entity ( $b$ ) contained in  $Assert_\beta(a,b)$  and  $\omega_K(r_i)$  be  $Rel(Assert_\beta(a,b))$ , multiplied by a temporal decay function  $\tau(\Delta t)$ . In the following section, we will show the benefit of augmenting a trust model with such a mechanism by presenting simulation results.

## 3 Simulation

For simulations, we have chosen to extend the trust model FIRE [14]. This particular model was chosen for its extensible nature, compatibility with the pro-

posed testing scenario and relative recency. Particularly its decentralized nature made it well-suited to our overall goal of mimicking the dynamics of reputation spreading in complex social networks.

### 3.1 Scenario

In order to assess the value of adding rumor-spreading mechanisms to trust models, we will evaluate the potential average gain an entity can expect when using a trust model with rumor-like information dissemination (FIRE w/ Rumor-Spreading), a non-augmented trust model (FIRE w/o Rumor-Spreading) and no trust model at all. For this, we will generally follow the methodology put forth in [14], with some changes to the scenario in order to better fit assumptions, for instance regarding trust dissemination [11].

An agent population consisting of consumer and provider agents is seeded in a spherical environment [14]. However, differing from the FIRE standard test methodology, the placement of consumer agents is influenced by an underlying complex social network [11], either a random graph [9] or a highly clustered acquaintance network [16]. Long-range connections between consumer agents that could not be placed together spatially are maintained in order to simulate the small-world effects of social networks [26]. During the simulation, recommendations and published opinions are communicated harnessing the underlying social network. Provider discovery and service delivery is handled in accordance with the neighborhood-based search employed in [14]. Provider selection, from a set of discovered providers, is also handled in accordance with the proposal from the FIRE testbed, using a standard Boltzman exploration strategy [18] in order to address the exploration-vs-exploitation dilemma.

Additionally, the testing methodology for assessing the advantage of agents equipped with a trust model and those without has largely been adopted from [14], as well. After selecting and interacting with a provider, the consumer gains or loses utility, dependent on the performance of the provider. Regarding this performance, the provider population is divided into three distinct sub-populations of consistent providers (*good*, *ordinary* and *bad*), as well as one of *intermittent* providers. Actual performance of providers is represented by a random variable, computed according to the sub-population a provider belongs to (for *bad*:  $\mu \in [-1, 0], \sigma = .2$ , for *ordinary*:  $\mu \in [0, .5], \sigma = .2$ , for *good*:  $\mu \in [.5, 1], \sigma = .1$ , according to a normal distribution; for *intermittent*: uniformly distributed in  $[-1, 1]$ ). A provider’s expected performance  $\mu$  is set at creation, its actual performance is determined per interaction. Time is measured in rounds, with events taking place during the same round occurring simultaneously. During each round, each agent chooses probabilistically, according to an individual activity level, whether it interacts with a provider. The utility score of every interaction is recorded by each agent, in order to assess the average utility gained or lost each round. Default experimental variable values were retained from the standard FIRE testbed. This includes: number of simulation rounds  $N = 500$ , number of provider agents  $N_P = 100$  (subdivided into good  $N_{PG} = 10$ , ordinary  $N_{PO} = 40$ , bad  $N_{PB} = 45$  and intermittent  $N_{PI} = 5$ ), number of consumer agents per test

group  $N_C = 500$ . Further default parameters, such as component coefficients and reliability function parameters were also retained, with the exception of the referral length threshold  $n_{RL}$ , which was set to permit only 1-hop referrals, as per the rationale in section 2.1.

The component coefficient for the proposed rumor-spreading component was set to  $W_{RS} = .5$ , identical to that of the component measuring direct witness information. As scaling function  $f$ , a generalized logistic function was applied, if the total number of informative votes on a published opinion was below 12, else it was set to 1. The temporal decay function was set to  $\tau(\Delta t) = \exp(-1 \cdot \Delta t / (-10 / \ln(0.5)))$ ,  $T_{\Delta Rel} = .1$  and  $p_{stifle} = .33$ . Published opinions on providers were issued probabilistically ( $p = .75$ ) if an agent gained utility in the top ten percentile range, has had five consistent interactions with a provider in the top twenty percentile range, or has lost utility from an interaction. Agents did not issue replacement opinions, unless the behavior of an agent was intermittent or decreased significantly. Furthermore, for rumor-like propagation, each agent received a send queue, into which the agent prioritized incoming published opinions for further propagation in future rounds, according to the following order: age, trust in the latest propagator and reliability of the message. Propagation was limited to 5 published opinions per round per propagating agent.

### 3.2 Results

In order to assess the benefit of augmenting a trust model with rumor-spreading, we investigated the average utility received per agent and round for three groups of agents. Agents choosing providers randomly, without support of any trust model, form the baseline for comparison. Against this baseline, two identical implementations of the FIRE trust model [14] were tested, one of these augmented with rumor-spreading. Simulations were run under both stable conditions and dynamic conditions, which included both provider and consumer population fluctuations, provider performance improvement/deterioration, and agent mobility, under the testbed default dynamic parameters presented in [14].

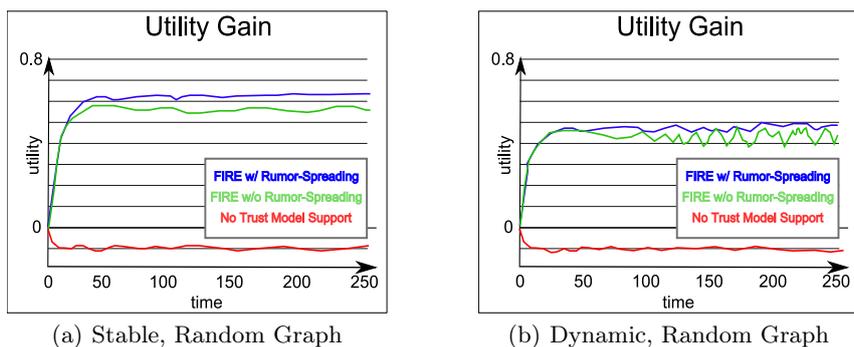


Fig. 2. Average utility gain under stable and dynamic conditions.

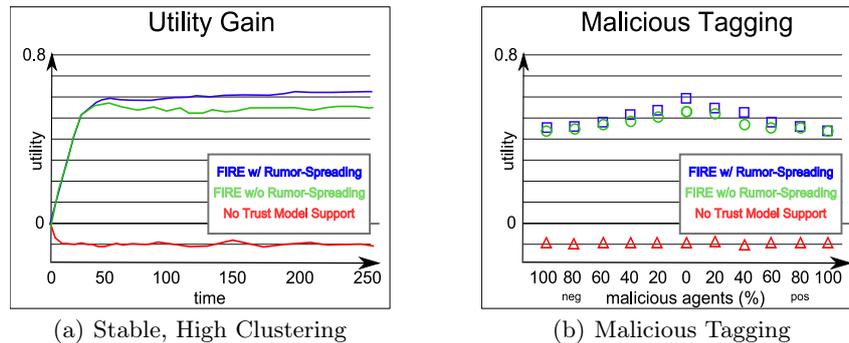
Trust-equipped agents consistently outperformed the no-trust group by a considerable margin. Under stable conditions, the agents augmented with a rumor-spreading component in turn outperformed the population equipped with a standard FIRE implementation by a margin of  $\approx .06$ , once a stable level of average utility was achieved after the first few interactions (cf. Fig. 2(a),3(a)). When comparing differently clustered social network structures used to communicate recommendations and published opinions, overall average utility changes were small (random graph (low clustering):  $\approx .63$  for rumor-spreading FIRE vs.  $\approx .57$  non-rumor-spreading FIRE; acquaintance graph (high clustering):  $\approx .61$  rumor-spreading FIRE vs.  $\approx .55$  non-rumor-spreading FIRE). High significance ( $p < 2.2e^{-16}$ ) was established by a signed-rank test [27]. However, in the highly clustered network, the number of initial rounds before a relatively constant utility level was reached, was almost double that of the random graph network. This behavior is congruent with findings regarding diffusion speed in [11].

In a simulation in which dynamic conditions introduce noise to the environment, the overall performance of FIRE, both with and without rumor-spreading augmentation, is lower than under stable conditions (cf. Fig. 2(b)). Agents with the additional component still maintain an advantage over agents equipped with pure FIRE, while both populations continue to perform considerably better than the no-trust population. Besides a slightly better performance, the rumor-spreading component can be seen to induce a stabilizing effect on the trust model. While performance of regular FIRE is subject to fluctuation, the change in average utility for the augmented model is much smoother.

In addition to being effective in reputation information dissemination, the proposed mechanism is also resilient to malicious tagging. This resilience can be attributed to the conservative reliability measure, based upon pairwise agreement computation, as well as its adaptability to the standard FIRE recommendation feedback mechanism, that effectively excludes unreliable recommenders from the recommendation process. Figure 3(b) displays the response to malicious tagging, as per the malicious tagging procedures described in [14]. Specifically, average utility gain over 200 time steps within the recommender population is plotted against the percentage of maliciously tagging recommenders. These agents either exaggerate or diminish a recommendee’s reputation by adding or subtracting a uniformly distributed random variable in  $[0.01, 0.2]$  to the utility gain they expect from the recommendee. An augmented trust framework can be clearly seen to result in higher average utility for those agents using it, as compared to agents equipped with a non-augmented implementation of the same framework.

## 4 Conclusions and Future Work

Simulation results indicate that augmenting trust models, as prototypically shown for the FIRE model [14], is beneficial to agent performance. The augmented model performed consistently better than the standard model by harnessing socially-filtered rumor-like information. The associated communication overhead



**Fig. 3.** Average utility gain under stable conditions in a highly clustered acquaintance network. Response to maliciously tagging agents.

for propagating rumor-like information can be partially mitigated by decoupling the provisioning process from the on-request model typically used in trust establishment. The voting mechanism, employing a standardized measure, displayed sufficient power to guarantee the reliability of the propagated published opinions. In the presented form, privacy issues of the protocol, which relies on agents issuing and forwarding published opinions but does not mandate them to do so, can be met through pseudonymization services and group key schemes. In order to harden the protocol against collusion attacks by malicious agent, tracking mechanisms and heuristics can be employed, for which a more sophisticated privacy protection mechanism would be required.

Both the exact structure of human acquaintance and trust networks in computational contexts, as well as the way that reputation information is communicated over them still leave considerable room for future investigations. Modeling these networks and applying computational procedures to them can not only serve to better understand human action, but also to assist online users, for instance by offering an automated, distributed (p2p) recommendation network. With the persistent popularity of social networking sites, the integration of socially augmented information systems is a logical next step in assuring reliable service in internet commerce.

## References

1. Abdul-Rahman, A., Hailes, S.: A distributed trust model. *New Security Paradigms* 97 (1997)
2. Agliari, E., Burioni, R., Cassi, D., Neri, F.M.: Efficiency of information spreading in a population of diffusing agents. *Phys. Rev. E* 73, 046138 (2006)
3. Castelfranchi, C., Falcone, R.: Principles of trust for mas: cognitive anatomy, social importance, and quantification. In: *Proceedings of the Third International Conference on Multi-Agent Systems* (1998)
4. Castelfranchi, C., Falcone, R.: *Trust and Deception in Virtual Societies*, chap. Social Trust: A cognitive approach, pp. 55–90. Kluwer Academic Publishers (2001)

5. Castellano, C., Fortunato, S., Loreto, V.: Statistical physics of social dynamics. *Reviews of Modern Physics* 81, 591–646 (2009)
6. Daley, D.J., Gani, J.: *Epidemic Modeling*. Cambridge University Press (2000)
7. Daley, D.J., Kendall, D.G.: Epidemics and rumors. *Nature* 204, 1118 (1964)
8. Dodds, P.S., Watts, D.J.: Universal behavior in a generalized model of contagion. *Physical Review Letters* 92 (21), 218701 (2004)
9. Erdős, P., Rényi, A.: On the evolution of random graphs. *Publications of the Mathematical Institut of the Hungarian Academy of Sciences* 5, 17–61 (1960)
10. Gambetta, D.: Can we trust trust? In: *Trust: Making and Breaking Cooperative Relations*. Basil Blackwell (1988)
11. Hauke, S., Pyka, M., Borschbach, M., Heider, D.: *Information Retrieval and Mining in Distributed Environments*, chap. Reputation-based Trust Diffusion in Complex Socio-Economic Networks. Springer-Verlag Berlin Heidelberg (2010), <in press>
12. Hauke, S., Pyka, M., Heider, D., Borschbach, M.: A reputation-based trust framework leveraging the dynamics of complex socio-economic networks for information dissemination. *Communications of SIWN* 7, 54–59 (2009)
13. Hayes, A.F., Krippendorff, K.: Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures* 1 (1), 77–89 (2007)
14. Huynh, T.D.: *Trust and Reputation in Open Multi-Agent Systems*. Ph.D. thesis, University of Southampton (2006)
15. Jackson, M.O., Rogers, B.W.: Search in the formation of large networks: How random are socially generated networks? Tech. Rep. 0503005, EconWPA (2005)
16. Jin, E.M., Girvan, M., Newman, M.E.J.: Structure of growing social networks. *Phys. Rev. E* 64(4), 046132 (Sep 2001)
17. Jøsang, A., Ismail, R., Boyd, C.: A survey of trust and reputation systems for online service provision. *Decision Support Systems* 43 (2), 618–644 (2007)
18. Kaelbling, L.P., Littman, M.L., Moore, A.W.: Reinforcement learning: A survey. *Journal of Artificial Intelligence Research* 4, 237–285 (1996)
19. Maki, D.P., Thompson, M.: *Mathematical Models and Applications, With Emphasis on the Social, Life and Management Sciences*. Prentice-Hall (1973)
20. Marsh, S.: *Formalising Trust as a Computational Concept*. Ph.D. thesis, Department of Computing Science and Mathematics, University of Stirling (1994)
21. Moreno, Y., Nekovee, M., Pacheco, A.: Dynamics of rumor spreading in complex networks. *Phys. Rev. E* 69, 066130 (2004)
22. Mui, L., Mohtashemi, M., Halberstadt, A.: A computational model of trust and reputation. In: *Proceedings of the 35th Hawaii International Conference on System Science*. pp. 280–287 (2002)
23. Sabater, J.: *Trust and Reputation for Agent Societies*. Ph.D. thesis, Universitat Autònoma de Barcelona (2003)
24. Sabater, J., Sierra, C.: Review on computational trust and reputation models. *Artificial Intelligence Review* 24, 33–60 (2005)
25. Wasserman, S., Faust, K.: *Social Network Analysis: Methods and Applications*. Cambridge University Press (1994)
26. Watts, D.: *Small Worlds*. Princeton University Press (1999)
27. Wilcoxon, F.: Individual comparisons by ranking methods. *Biometrics* 1, 80–83 (1945)
28. Yu, B., Singh, M.: Towards a probabilistic model of distributed reputation. In: *Proceedings of the Fourth Workshop on Deception, Fraud and Trust in Agent Societies*, Montreal (2001)
29. Yu, B., Singh, M.: Distributed reputation management for electronic commerce. *Computational Intelligence* 18 (4), 535–549 (2002)