The complex facets of reputation and trust

Karl Aberer[1], Zoran Despotovic[2] , Wojciech Galuba[1], Wolfgang Kellerer[2]

[1] Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland
{karl.aberer,wojciech.galuba}@epfl.ch
[2] DoCoMo Communications Laboratories Europe, Munich, Germany
{despotovic,kellerer}@docomolab-euro.com

Summary. Trust and reputation systems have proven to be essential to enforcing cooperative behavior in peer-to-peer networks. We briefly describe the current approaches to building reputation systems: social networks formation, probabilistic estimation and game theoretic models. We then observe that all of the current models make a number of simplifying assumptions that may not necessarily hold in real networks, such as either irrational (probabilistic) or completely rational behavior, instant propagation of reputation information and homogeneity of interactions. We argue that dropping those assumptions and allowing more degrees of freedom is necessary in order to construct more realistic and rich reputation models. We support our argument by citing reputation research done in economics, evolutionary psychology, biology and sociology and and consider models that take into account adaptive behavior changes, co-evolution of behaviors, bounded rationality and variable interaction patterns. We then outline how those complexities can be dealt with and point out main directions for the future study of more realistic and less constrained reputation models that can potentially lead to construction of more secure, responsive and cooperative peer-to-peer systems.

1 Introduction

Reputation systems have proven to be essential to enforcing cooperative behavior in peer-to-peer networks. Many solutions have been proposed [19, 32, 23, 14, 7, 3], each employing a different model of computing trust, disseminating and storing reputation data and responding to non-cooperation in the network [26]. In this paper we focus on the reputation and trust models themselves rather than practical considerations of implementing and deploying a reputation system. We begin with the description of the basic concepts, then survey the current approaches, examine the different assumptions commonly made by the different reputation and trust models and propose ways in which they can be relaxed or extended.

2 Fundamentals

Assume a set of *nodes* continuously engaging in bilateral *interactions*. For simplicity we assume that a single interaction always involves a pair of nodes and that a interactions involving a larger group of nodes can always be decomposed into a set of binary interactions.

Each interaction has an associated *benefit* and *cost*. These two values are normally such that nodes face a Prisoner's Dilemma (PD) [13]. It is beneficial for the node to cooperate only if the other node cooperates as well, otherwise it is better to defect.

When Alice interacts with Bob it can gain more if it is able to predict that Bob will cooperate. The extent to which a node believes the other will cooperate is the extent to which a node *trusts* the other node. There are a number of ways this belief can be inferred and they are captured by the different *trust models*. One of the inferences that can be made is: if Alice cooperated with Bob then it implies Alice will also cooperate with Carol. If this inference is applied universally, the collective actions of Alice form a commonly shared belief among the other nodes of how likely Alice is to cooperate. This belief is what is termed *reputation*. In the paper we will focus on reputation-based models of trust, whose computation solely depends on the actions of the peers instead of relying on other elements such as third party guarantors of trust (e.g. PKI) or virtual currency for which trust can be purchased etc.

3 State-of-the-art

In reputation-based models trust towards a given node A is determined based on the past actions of A. Every node V_i only has information about the actions of A that V_i itself experienced. To compute the reputation of A, nodes need to exchange the information about the actions of A that they have observed. This exchange and the subsequent computation of reputation can proceed in many ways.

There are four classes of approaches [8]: social networks, probabilistic estimation, game-theoretic models and evolutionary approaches.

3.1 Social networks

The social network approach assumes an existence of a digraph of social links between nodes. The interactions between the nodes proceed along the links and each link has a trust value associated with it. That value is updated based on the interactions between the nodes at the two ends of the link. A node V can compute the trust value for another non-neighbor node W by aggregating trust values from other nodes in the following way:

1. enumerate (all) paths from W to V
2. aggregate trust values along the paths
3. merge the results of aggregation at V as the final trust value

The social network approaches vary in the details of the three above steps: what domain is used to represent trust, what the selected paths are, what are the aggregation and merging functions. Trust values are either computed on demand between specific W and V or simultaneously for all nodes using some form of iterative methods that converge on an eigenvector of trust values.

3.2 Probabilistic estimation

The computations in social networks produce trust values that are hard to interpret. In particular, given a trust value for the node A it is hard to translate that value into the probability that A will cooperate. But this can be rectified if the assumption about probabilistic behavior of the nodes is made explicit and then well known probabilistic estimation techniques such as Bayesian estimation and maximum likelihood estimation are used to compute the trust of a given peer [9]. This is what probabilistic estimation methods do. As an example, consider a network consisting of peers having associated innate probabilities of performing trustworthy. Denote by θ_j the probability of peer j. Assume that peer j interacted with n other peers $p_1, \ldots, p_n$ and its performances in these interactions were $x_1, \ldots, x_n$, where $x_i \in \{0, 1\}$ (1 denoting the honest performance and 0 the dishonest one). When asked to report on peer j's performances witnesses $p_1, p_2, \ldots, p_n$ may lie and misreport. Assuming that they lie with specific probabilities, say l_k for peer p_k, the probability of observing report y_k from peer p_k can be calculated as:

$$P[Y_k = y_k] = \begin{cases} l_k(1-\theta_j) + (1-l_k)\theta_j \text{ if } y_k = 1 \\ l_k\theta_j + (1-l_k)(1-\theta_j) \text{ if } y_k = 0. \end{cases} \quad (1)$$

By definition, the likelihood function associated with a random sample of reports $y_1, y_2, \ldots, y_n$ is:

$$L(\theta_j) = P[Y_1 = y_1]P[Y_2 = y_2] \cdots P[Y_n = y_n]. \quad (2)$$

After collecting the reports on the peer it is about to interact with, the trust computing peer just has to make this product and find θ_j that maximizes it. This number is the maximum likelihood estimate of the unknown probability. To do this, the computing peer must have good estimates of the parameters $l_1, \ldots, l_n$. They can be made by comparing own performances with reports on them. Note also that the own experiences are seamlessly integrated into this model - the trust computing source peer i just has to put $p_i = 1$ for his own experiences x_i. As another advantage of the probabilistic methods, we emphasize that, when compared to social networks, they bring a substantial reduction of the communication overhead. The reason is that they deal only with feedback on the target peer, while social networks essentially aggregate all available feedback, i.e. opinions of all peers about all other peers.

3.3 Game theoretic approach

In game theoretic approaches to the reputation systems it is often assumed that the players are perfectly rational in the sense that they are only interested in maximizing their own payoffs. These assumptions allow the computation of Nash equilibria as strategy profiles where peers have no incentive to deviate. Normally, game theoretic modeling of reputation effects requires repeated interaction and uncertainties among the players with respect to their opponents' payoffs [22]. More recently, there have been attempts to extend these models in order to more closely model real world settings. Most notably, the two important models are: private and public monitoring games. In these games players do not observe each other's actions but only their signals. In private monitoring games [21] the signals are different for different players, while in public monitoring games [24], all peers observe the same signals about the actions of other peers.

However, we see a number of problems with respect to the application of game theoretic reputation models. One is related to the behavior. There are plenty of settings where the full rationality of the players cannot be expected. Any setting with human players would be an example. The second is the difficulty of introducing the rationality assumption into the reputation mechanism implementation itself.

3.4 Evolutionary approach

Game theorists have also approached the problem of cooperation in a population of PD-players from a more experimental angle. Most notably, Axelrod [4] has demonstrated the success of the tit-for-tat strategy in an Evolutionary Prisonner's Dillemma setting. In this setting pairs of players are involved in repeated PD games. Each player maintains a score, which is updated after every game round according to the PD payoff matrix. The players with the highest score are considered most fit and their strategies are replicated replacing other unfit strategies. The winning tit-for-tat strategy follows three simple rules:

1. *initially cooperate* - when interacting with an opponent for the first time, always cooperate
2. *punish* - if the opponent defected in the previous round, punish him by defecting
3. *forgive* - if the opponent cooperated in the previous round, cooperate even if there is a history of opponent's defection

The tit-for-tat strategy has been shown to be evolutionary stable, being able to drive into extinction small populations of invading defectors, that try to exploit cooperators. At the same time groups of tit-for-taters are always cooperating with each other, which allows them to accumulate score surplus which in turn can be used to fight against transient groups of defectors.

To be successful, the tit-for-tat strategy needs a setting in which the PD interactions are repeated many times for the same pair of players, which allows punishment to occur. In a large population of infrequently interacting individuals this may not be possible (e.g. eBay and its transactions). This observation led to the definition of a new setting in which every pair of players can only play one round of PD and never meet again. Building cooperation in this setting relies on the rule: "If A cooperates with B then B can reciprocate and cooperate with some other player C". This rule is termed *indirect reciprocity*, as opposed to the *direct reciprocity* rule followed by tit-for-tat. In this case, to build cooperation players can no longer rely on private observation of the actions of the opponent. Once an observation is made, remembering that observation is pointless since all interactions are one-shot and such observation can never be used to make cooperation decisions. Hence there arises the need to exchange observations with other players. This can be implemented by associating a public label with each player. All players can read the label, and all players except the owner of the label are allowed to change it. It has been shown that to enable sustainable cooperation only two states of the label are sufficient [20]. The two states correspond to good and bad reputation. When a pair of players interacts, their labels are modified according to their actions. The behavior of the player can be succinctly described as two functions: the *action function* and the *assessment function*. The action function takes the label of self and the opponent and produces the decision to either cooperate or defect. The assessment function is executed after the actions of both agents have taken place. The assessment function takes the label of self, the label of the opponent and the action of the opponent and produces the new value for the opponent's label. Since the outputs of the functions are binary, there is a relatively small number of all possible functions. There are exactly 16 possible action functions and 256 possible assessment functions, which together results in 4096 possible behaviors. Ohstuka et. al. [15] have performed a systematic experimental study of all those 4096 behaviors. Out of these they have found 8 evolutionary stable cooperative strategies, termed "the leading eight" (Table 1).

A population of agents using one of these strategies is able to sustain cooperation and drive out of existence any small population of defectors and/or reputation liars (i.e. players that set the labels to "bad" value even though their opponent cooperated).

There is a remarkable similarity between tit-for-tat and the leading eight strategies. The leading eight strategies exhibit all the properties of tit-for-tat: initial cooperation, forgiveness and punishment for defection. Tit-for-tat can be implemented with one bit of local state in the player, leading eight strategies make this state public by storing it in the player's label.

assessment function:

	GG	BG	GB	BB
C	G	*	G	*
D	B	G	B	*

action function:

GG	GB	BG	BB
C	D	C	*

Table 1. The "leading eight" behaviors in the evolutionary indirect reciprocity game. G and B stand for good and bad reputation labels respectively. C and D stand for cooperation and defection. The GG, GB, BG, BB encode 4 possible states of the labels. The first letter is the label of self and the second letter is the label of the opponent. The 3 asterisks in the fields of the assessment function can take any value, hence 8 possible assessment functions are possible. The value at the asterisk in the action function is uniquely determined based on the choice of one of the eight assessment functions (for details refer to [15]). As can be seen from the tables, the leading eight behaviors are similar to tit-for-tat, bad behavior is forgiven after it is punished. In addition to that, punishment of bad behavior is justified, a good player defecting with a bad player is assessed as good.

4 Propagation of reputation information

If we compare the two cases - direct reciprocity and indirect reciprocity - they are two extremes in reputation information propagation. In the case of direct reciprocity it is sufficient to rely on privately gathered history of interactions with players, no propagation of reputation is necessary. On the other hand, in the case of indirect reciprocity, once two players interact, their reputation labels are updated and immediately available to all other players, the reputation information propagates instantaneously. When a reputation system is implemented in a peer-to-peer setting the assumptions about the propagation of reputation no longer hold. The character of reputation propagation is determined by the implementation. The question that arises is whether the non-instantaneous reputation propagation influences the performance of the reputation system. There is at least one piece of evidence [5] which suggests that delaying the communication of reputation in games with imperfect private monitoring leads to more efficient equilibria. Taking the propagation of reputation information into account might lead to the discovery of entirely new phenomena.

Question 1 *How do the reputation propagation dynamics influence the performance of the reputation system?*

The propagation of reputation can not only be delayed, it may also be possible to propagate it partially while still maintaining the reputation system performance at an acceptable level. Participating in a reputation system incurs a cost to the peers, the smaller the fraction of nodes that need to participate in each reputation update the smaller the load on the system. We have performed simulations to test the impact of limited reputation information propagation

on the performance of the system (Figure 1). Experiments indicate that it is

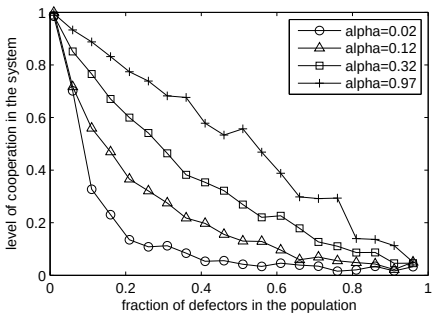

Fig. 1. A population of honest agents using one of the leading eight strategies (see Section 3.4) is pitted against a population of defectors, who always defect and propagate negative reputation information about others. We vary the number of defectors in the population and observe the level of cooperation in the system measured by the fraction of interactions amongst the honest agents in which both agents cooperate. Every reputation update is propagated to a fraction α of the whole population chosen uniformly at random. We repeat the experiment for different values of α. We can observe that if the reputation updates are propagated to a few agents only (2%) even a small number of badmouthing defectors can subvert cooperation. On the other hand, the reputation propagation rate set to 30% is sufficient to allow practically linear graceful decrease in cooperation level as the number of defectors increases.

sufficient to make the reputation information available to 30% of the agents to obtain performance that is close to the performance of the system with full propagation. This suggests that there are substantial communication savings to be gained by simply limiting the propagation of the reputation information.

Question 2 *Is it necessary to propagate the reputation information to all the nodes to have a robust reputation system?*

Question 3 *How does the fraction of nodes to which reputation information is propagated influence the performance of the system?*

Question 4 *How to choose the fraction of nodes to which the reputation information is propagated?*

5 Bounded rationality

Game theorists have considered imperfect monitoring games in which noise is allowed to occur in the system: imperfect observation of other players' actions, imperfect action execution, error-prone reputation information exchange, etc.

While considering limitations of perception of the players, game theory still usually assumes that the players are absolutely rational. However, they may have limited resources available to them to compute their behavior. Nash equilibria have been shown to be NP-hard to compute [12]. In the extreme, being unable to compute their behavior peers can behave entirely irrationally (randomly). This set of limitations is commonly termed *bounded rationality*. Conlisk [6] provides a plethora of empirical evidence from economics and experimental psychology in support of bounded rationality. The key observations are that:

- bounded rationality can explain a number of empirical anomalies in economics for which unbounded rationality models fail
- rationality is scarce, good decisions are costly, they require both reliable information, which is difficult to obtain and computational power
- bounded rationally leads people to imitate behaviors of others, which is cheaper than computing the behavior on their own

Given the predictive success of bounded rationality models, questions arise:

Question 5 *How can we incorporate bounded rationality into reputation models?*

Question 6 *What are the bounds on rationality in peer-to-peer systems and how can they influence the dynamics of cooperation and reputation?*

6 Behavioral evolution

In the previous section we have already mentioned how imitation plays a role in selection of behaviors by agents. When a behavior is replicated its utility is locally evaluated by the agent. If the utility of the behavior is low it is promptly replaced by another behavior. This creates an evolutionary setting in which behaviors are replicated by imitation and selected by the agents for utility. An agent might use a set of behaviors (rules of behavior) and each of them can be individually imitated, creating a setting in which groups of mutually dependent behaviors co-evolve. The two main mechanisms of behavioral imitation in human societies are: payoff-biased transmission - imitating the behavior of the most successful individuals and conformist transmission - imitating the most frequent behavior [16].

How can we relate the above facts about behavioral evolution to interacting populations of selfish peers in peer-to-peer systems? First, we must clarify that it is not the peers that are selfish, but the human users of the peer-to-peer software. It is the users themselves who decide how their peers should behave. Hence, we could conjecture that a lot of the social mechanisms described above are driving the evolution of peer behaviors. This conjecture is confirmed by the following empirical evidence. A peer-to-peer file sharing software called

eMule [2] is open source, which allows anyone to make modifications to it and distribute them. This has given rise to a number of mutated versions of the base eMule client, the so-called "mods" [1]. There are mods that protect the user privacy by encrypting downloaded data, there are mods which implement various bandwidth saving heuristics, there are extremely non cooperative mods that cut off uploads to other peers to conserve bandwidth, there are mods that detect non-cooperative mods and disconnect from them, there are even mods that detect those policing mods and use stealth techniques to hide their defection etc. These mods are constantly created and propagated via numerous websites and evaluated by users on various electronic forums. The social network of peer-to-peer system users selecting behaviors for their peers is tightly interrelated with the overlay network providing an arena for the execution of those behaviors selected by the users. Up to our knowledge there have been no attempts to study these two networks as one entity with all their dependencies.

Question 7 *How can we model behavioral evolution in peer-to-peer systems?*

Question 8 *How can we model the peer-to-peer software choices and modifications made by humans and how do they affect the performance of the system?*

7 Second-order defection problem

In an indirect reciprocity setting with cooperation being sustained by the means of reputation there exists a following problem: in order for the reputation system to work, agents need to cooperate on exchanging reputation information and the information about the actions of other agents they have observed. Moreover, for the reputation system to be effective agents need to punish defectors which incurs additional costs. This creates a second-order cooperation problem, which could be solved by adding yet another reputation system on top of the existing one, but this in turn would lead to a third-order cooperation problem.

In peer-to-peer reputation systems research the problem is rarely explicitly addressed, the usual practice is to test the robustness of the system by introducing subpopulations of second-order defectors, i.e. peers that withhold or provide false reputation information. These evaluations only show that first-order cooperation can be sustained under a second-order defector invasion but it does not show that second-order cooperation is sustainable.

One of the game theoretic solutions to this problem is the construction of an providing incentives [18] to motivate agents to share their reputation information truthfully, however the solution relies on a third party to handle the payments. This and many other similar approaches simply reformulate the problem of second-order defection and delegate it to another, normally centralized system component. Up to now there has been no self-contained

distributed reputation system proposed that is free from second-order defection problem.

There is, however, a natural system that appears to have solved that problem - human society. Biologists and psychologists studying indirect reciprocity among humans have been trying to find the exact reasons for the remarkable stability of reputation and how it evolved [31, 27]. Many hypotheses have been proposed, most notably:

- group selection - Boyd et al. [28] suggest that cooperation can evolve by natural selection at the level of groups. Those groups that use reputation are more cooperative and hence more fit.
- conformist transmission - Heinrich et al. [16] show how weak conformity in populations can lead to the stabilization of reputation exchange and cooperation.
- costly signalling - Gintis et al. [30] show how using costly signals agents can advertise their quality as cooperators and in this way increase their reproductive success.

These mechanisms could be implemented and studied in artificial reputation systems potentially leading to increased performance and stability of second-order cooperation.

Question 9 *How can we apply the known reputation stability mechanisms from natural systems to engineering peer-to-peer systems free from second-order defection problem?*

8 Inhomogeneous interactions

In models of reputation systems it is frequently assumed that the structure of interactions between agents is homogenous, i.e. each agent is equally likely to interact with any other agent. This assumption allows the construction of tractable analytical models, however, in practice the pattern of interactions in the system may not be homogenous, which may produce large deviations from the predictions of the models. For example nodes that interact with a large number of other nodes may need to rely more on reputation information exchange and nodes that frequently interact with a small subset of nodes may rely more on bilateral tit-for-tat strategies and may have no incentive to share the reputation. These two types might need to co-exist in the same network. More complex behaviors are possible, a group of nodes that are highly interacting with each other may choose to collude by artificially increasing each other's reputations but defecting with other nodes that are not part of the group. Once non-homogeneous interactions are allowed there is no single wining behavior, such as Ohtsuka's leading eight or Axelrod's tit-for-tat. A complex set of mutually dependent behaviors can successfully coexist.

In overlay routing substrates the structure of interactions is normally determined by the underlying overlay maintenance algorithm - the interactions are packets forwarded by the nodes to their neighbors. In the case of DHTs the interactions are the key access and insertion requests, which are determined by the particular data placement strategy, normally a hash function. The inhomogeneities in the structure of interactions in any of those cases may warrant the existence of different equilibrium behaviors for different nodes.

Question 10 *What is the character of interaction inhomogeneities in peer-to-peer networks?*

Question 11 *How can those inhomogeneities influence the behavior of the selfish peers exchanging reputation information?*

So far, we have assumed that inhomogeneities arise from some external mechanism outside the peer's control. In general, however, a peer might decide what peers it interacts with based on its selfish choice. For example, a peer may choose to interact less frequently with low reputation peers. Selection of who to interact with becomes part of the peer behavior, which leads to a recursive problem: the structure of interactions determines the optimal behaviors at every node and the behaviors of nodes determine the structure of interactions.

A number of studies have looked at network formation by selfish peers [29, 11, 17]. However, all of the studies assume behavioral homogeneity of peers, i.e. all peers having the same utility function. Also, none of the studies consider both network formation and cooperation building via reputation as a single problem.

Question 12 *How can peers use the reputation information to choose what peers they want to interact with? What is the structure and dynamics of the resulting interaction network?*

9 Identity stability

Most reputation systems rely on the assumptions that identities of the agents are stable and can be reliably used, however, in contrast to human societies, identities in a peer-to-peer systems are low cost and easy to change. A malicious peer whose reputation is low, can leave the system and rejoin under a different identity thus clearing the whole history of its defections. A malicious peer can also assume a number of identities to have significant presence in the network [10]. Identity can also be stolen to take advantage of the reputation of previous owner.

A well-known solution to the problem of identities is public key infrastructure. However, maintaining a hierarchy of trusted third parties creates scalability problems as well as introducing a single point of failure. Another widely employed solution is increasing the cost of identities by initializing the

reputation of newly coming peers to a low value and making the peers gradually build their reputation. This, however, creates a disadvantage for short lived peers who loose their identity every time they depart from the system and during their short lifetime are not able to accumulate enough reputation to gain any benefit from participation in the system.

When considering identity, researchers commonly assume one of the two extremes: either a cheap, easy to change identities or expensive, reliable ones. However, there exist cases which lie in between. For example, when two peers open a TCP connection to communicate through it, the stability of the identities at both ends of the TCP link is guaranteed. This concept of pairwise identity stability can be extended to arbitrary groups of communicating peers within which peer identities are stable, identity stability needs to be associated with a particular scope.

We may also add assumptions about partial perception of identity, i.e. a peer might only be able to determine that a node belongs to some larger group but not pinpoint exactly which node it is. For example, a node might be identified as belonging to a university campus, but the individual identity of the peer might be unknown. This creates new challenges and adds more complexity to the already wide range of possible behaviors in a reputation system. Up to our knowledge, partial perception of identities has not been considered in the context of reputation systems.

Question 13 *What are the minimal assumptions on the stability and perception of identity needed to construct a robust reputation system?*

Identity is inextricably linked with anonymity and privacy in peer-to-peer networks. Having accurate identity models might enable the designers to make more precise statements about the anonymity guarantees in their peer-to-peer systems [25].

Question 14 *Can cooperation be sustained while maintaining anonymity in a peer-to-peer system? What are the tradeoffs?*

10 Conclusions

Each of state-of-the-art approaches to reputation systems for peer-to-peer networks is based on a set of assumptions about the target deployment environment. We have demonstrated how breaking of some of these fundamental assumptions leads to unexpected phenomena and complex peer behavior. Clearly, there exists no single universal solution that can work well in all distributed environments, instead the properties of the environment should be precisely determined before designing a reputation system. We have identified the main dimensions along which these environment properties can be categorized:

- **communication model** - how information propagates in the environment, how costly the propagation is, this influences the speed at which reputation information can be disseminated and how many peers it may reach
- **computational constraints** - how costly computation and local storage are, these assumptions determine the degree to which peers' rationality is bounded
- **peer software dynamics** - how selfish users deploy new software, how software is modified, these processes drive the behavioral evolution of the system and put constraints on how fast new behaviors can be deployed or enforced in an existing system
- **interaction model** - how peers interact, to what degree they can choose their interaction partners, these properties of the system can strongly influence the reputation dynamics and the choice of the optimal behavior
- **identity model** - whether identity might change, how identity is created and represented and to what degree it can be accessed by other peers, this determines the level of privacy and anonymity and the precision at which statements about the reputation of individual peers can be made
- **peer goal dynamics** - what the goals of the peers are, how they change over time, this describes the behavioral heterogeneity of the population and at the same time groups of peers with malicious goals can be used to model many forms of attacks on the system

All of these environment properties strongly influence the design choices that need to be made when constructing a reputation system. How do the environment properties constrain the performance of the reputation system? What are the combinations of environment properties that fundamentally prevent from building any cooperation in the system? What is the best formal model of the distributed target environment which allows to make precise statements about all of its properties? These and many other problems constitute a new and exciting agenda for trust and reputation research in peer-to-peer systems.

References

[1] emule-mods `http://www.emule-mods.de/`, 2006. [Online; accessed 22-Feb-2006].

[2] emule, open source p2p file-sharing client, `http://www.emule-project.net/`, 2006. [Online; accessed 22-Feb-2006].

[3] Karl Aberer and Zoran Despotovic. Managing trust in a peer-2-peer information system. In *CIKM*, pages 310–317, 2001.

[4] Robert M. Axelrod. *The Evolution of Cooperation.* Basic Books, 1984.

[5] Olivier Compte. Communication in repeated games with imperfect private monitoring. *Econometrica*, 66(3):597–626, May 1998.

[6] John Conlisk. Why bounded rationality? *Journal of Economic Literature*, 34(2):669–700, June 1996.

[7] Ernesto Damiani, Sabrina De Capitani di Vimercati, Stefano Paraboschi, and Pierangela Samarati. Managing and sharing servents' reputations in P2P systems. *IEEE Trans. Knowl. Data Eng*, 15(4):840–854, 2003.

[8] Zoran Despotovic. *Building Trust-Aware P2P Systems: From Trust and Reputation Management to Decentralized e-Commerce Applications.* Diploma thesis, EPFL, Swiss Federal Institute of Technology, Lausanne, 2005.

[9] Zoran Despotovic and Karl Aberer. A probabilistic approach to predict peers' performance in p2p networks. In *Eighth International Workshop on Cooperative Information Agents, CIA 2004*, Erfurt, Germany, 2004.

[10] Douceur. The sybil attack. In *International Workshop on Peer-to-Peer Systems (IPTPS), LNCS*, volume 1, 2002.

[11] Fabrikant, Luthra, Maneva, Papadimitriou, and Shenker. On a network creation game. In *PODC: 22th ACM SIGACT-SIGOPS Symposium on Principles of Distributed Computing*, 2003.

[12] Fabrikant, Papadimitriou, and Talwar. The complexity of pure nash equilibria [extended abstract]. In *STOC: ACM Symposium on Theory of Computing (STOC)*, 2004.

[13] D. Fudenberg and J. Tirole. *Game Theory.* MIT Press, Cambridge, MA, 1991.

[14] Minaxi Gupta, Paul Judge, and Mostafa H. Ammar. A reputation system for peer-to-peer networks. In *NOSSDAV*, pages 144–152, 2003.

[15] Y. Iwasa H. Ohtsuki. How should we define goodness? reputation dynamics in indirect reciprocity. *Journal of Theoretical Biology*, 231:107–120, 2004.

[16] Joseph Heinrich and Robert Boyd. Why people punish defectors. *J Theor Biol.*, 208(1):79–89, Jan 2001.

[17] Matthew O. Jackson and Asher Wolinsky. A strategic model of social and economic networks. *Journal of Economic Theory*, 71(1):44–74, 1996.

[18] Radu Jurca and Boi Faltings. Towards incentive-compatible reputation management. In L. Korba R. Falcone, S. Barber and M. Singh, editors, *Trust, Reputation and Security: Theories and Practice*, volume Lecture Notes in AI 2631, pages 138–147. Springer-Verlag, Berlin Heidelberg, 2003.

[19] Sepandar D. Kamvar, Mario T. Schlosser, and Hector Garcia-Molina. The eigentrust algorithm for reputation management in P2P networks. In *WWW*, pages 640–651, 2003.

[20] Michihiro Kandori. Social norms and community enforcement. *Review of Economic Studies*, 59(1):63–80, 1992.

[21] Michihiro Kandori and Hitoshi Matsushima. Private observation, communication and collusion. *Econometrica*, 66(3):627–652, May 1998.

[22] David M. Kreps and Robert Wilson. Reputation and imperfect information. *Journal of Economic Theory*, 27:253–279, 1982.

[23] Seungjoon Lee, Rob Sherwood, and Samrat Bhattacharjee. Cooperative peer groups in NICE. In *INFOCOM*, 2003.

[24] George J. Mailath and Stephen Morris. Repeated games with almost-public monitoring. CARESS Working Papres almost-pub, University of Pennsylvania Center for Analytic Research and Economics in the Social Sciences, August 1999.
[25] Sergio Marti and Hector Garcia-Molina. Identity crisis: Anonymity vs. reputation in P2P systems. In *Peer-to-Peer Computing*, pages 134–141, 2003.
[26] Sergio Marti and Hector Garcia-Molina. Taxonomy of trust: Categorizing P2P reputation systems. *Computer Networks*, 50(4):472–484, 2006.
[27] M. A. Nowak and K. Sigmund. Evolution of indirect reciprocity. *Nature*, 437:1291–1298, 2005.
[28] Samuel Bowles Robert Boyd, Herbert Gintis and Peter J. Richerson. The evolution of altruistic punishment. *Proceedings of the National Academy of Sciences*, 100(6):3531–3535, Mar 2003.
[29] Tim Roughgarden. *Selfish Routing and the Price of Anarchy*. MIT Press, Cambridge, MA, 2005.
[30] Eric Alden Smith, Samuel Bowles, and Herbert Gintis. Costly signaling and cooperation. Working Papers 00-12-071, Santa Fe Institute, December 2000.
[31] Robert L. Trivers. The evolution of reciprocal altruism. *The Quarterly Review of Biology*, 46(1):35–57, Mar 1971.
[32] Li Xiong and Ling Liu. Peertrust: Supporting reputation-based trust for peer-to-peer electronic communities. *IEEE Trans. Knowl. Data Eng*, 16(7):843–857, 2004.