

Reputation Dynamics and Convergence: A Basis for Evaluating Reputation Systems

Christopher J. Hazard and Munindar P. Singh

Department of Computer Science
North Carolina State University
Raleigh, NC 27695-8206, USA

{cjhazard,singh}@ncsu.edu

November 2009

Abstract

Reputation is a crucial concept in dynamic multiagent environments and not surprisingly has received considerable attention from researchers. However, despite the large body of related work on reputation systems, no metrics exist to directly and quantitatively evaluate and compare them. We present a common conceptual interface for reputation systems and a set of four *measurable* desiderata that are broadly applicable across multiple domains. These desiderata employ concepts from dynamic systems theory to measure how a reputation system reacts to a strategic agent attempting to maximize its own utility. We study a diverse set of well-known reputation models from the literature in a moral hazard setting and identify a rich variety of characteristics that they support. We discuss the implications, strengths, and limitations of our desiderata.

1 Introduction

An agent's reputation is the aggregation of publicly available information about the agent. Such information is not necessarily accurate. Trust and reputation are often used in a complementary fashion: an agent expects positive outcomes when interacting with another agent that has a reputation for being trustworthy. Some systems are best described as trust systems because therein agents determine whether another agent will do what it says it will, whereas others are best described as reputation systems because therein agents determine and propagate their beliefs about other agents. The mechanics of the two kinds of systems exhibit considerable overlap [Ramchurn et al., 2004].

Reputation is an important concept and computational reputation systems are popular primarily because there are strong intuitive connections between an agent's reputation and both the utility that it obtains and the utility another agent obtains when interacting with it. For example, an agent can obtain more money for the same products on eBay (<http://ebay.com>) simply by having a more positive reputation [Houser and Wooders, 2006]. A rational agent would only build and maintain a positive reputation if doing so maximizes utility. For example, in commerce environments, an agent can strategically build up and then expend its reputation in order to monopolize a market [Sen and Banerjee, 2006].

Many authors propose desiderata to motivate their trust and reputation systems [Huynh et al., 2006; Kamvar et al., 2003; Zacharia and Maes, 2000]. However, we are unaware of a general characterization

of desiderata for reputation systems that are quantitative, objective, and applicable across a wide range of domains. We present four desiderata, focusing on what quantitative properties make one reputation system more effective than another. Devising widely applicable metrics for trust is considered an important open problem [Barber et al., 2003] and is the focus of this work.

The primary purpose of a reputation system is to handle cases of *adverse selection* and *moral hazard* [Dellarocas, 2005]. Adverse selection occurs when agents have limited ability to change, for example, if a peer on a file sharing network supports limited upload bandwidth. In this case, other agents want to learn which agents have favorable attributes (significant upload bandwidth) so that they can choose agents with whom to interact. Moral hazard arises when one agent must reduce its utility in order to increase another's utility. An example of moral hazard is when one agent buys an item expecting it to be at or above a certain quality, but cannot measure the quality until after the purchase. Here, the seller would face the moral hazard of producing a lower quality item to reduce its costs.

When dealing with rational agents in a pure moral hazard setting, the game theoretic approach is to devise a folk theorem, possibly modifying the model to achieve desired equilibria. The analogous approach when dealing with pure adverse selection is to use probability and statistics to determine agents' types. Although these approaches are powerful in pure scenarios, most real-world applications do not cleanly fall into one of the pure scenarios.

Reputation is only meaningful if it can change over time to increase predictive accuracy in cases of adverse selection and to incentivize agents to cooperate in cases of moral hazard. Therefore, we approach reputation from a dynamic systems perspective. As our primary contribution, we motivate and formalize the following quantifiable desiderata.

Monotonicity. Agents who would provide favorable interactions should acquire better reputations than agents who would provide less favorable interactions. For example, a seller who always offers high-quality items at a low price should have a better reputation than an agent who produces defective items that it advertises as being of high-quality (and thus sells at a high price).

Accuracy. Reputation measurements should be accurate regardless of prior beliefs. For example, if a buyer incorrectly believes that a seller produces high-quality items, the buyer should quickly learn an accurate reputation value for the seller.

Convergence. Agents' reputations should converge quickly. For example, it is preferable to be able to learn after a smaller number (rather than a greater number) of interactions whether a seller offers high or low-quality products, regardless of past beliefs, provided the seller keeps to its type.

Unambiguity. An agent's reputation should be asymptotically unambiguous, meaning an agent's asymptotic reputation should be independent of any a priori beliefs about the agent held by some observing agent. An unambiguous reputation system would, as the number of interactions tends toward infinity, always yield the same reputation for a given agent regardless of the specific interactions. Consider two otherwise identical buyers (identical in their valuations for goods of a given quality, utility functions, capabilities, influence over peers, and so on) initially disagreeing about a seller's reputation. Both buyers should converge to an agreement about the seller's reputation after a sufficiently large number of interactions, assuming the seller behaves steadily in the same manner with each buyer.

Our desiderata apply to both adverse selection and moral hazard, with or without the propagation and aggregation of reputation information. The measurements from the desiderata can answer a wide range of questions, such as whether agents would benefit from using a specific reputation system, how stable the

system is, and how quickly agents can build up or lose their reputation. Rather than examine and compare reputation systems against a list of possible attacks [Huynh et al., 2006; Kerr and Cohen, 2009; Kamvar et al., 2003], we look at general dynamical properties of the system as affected by strategic agents. Our desiderata are useful across many types of reputation systems, regardless of whether the reputation system combines moral hazard with adverse selection, involves interactions in less clearly defined environments, or how difficult it is to solve analytically.

Throughout this paper, we distinguish two roles that an agent plays in a reputation system. An agent is a *rater* when evaluating others and is a *target* when it is being evaluated. An agent may take on both roles of target and rater simultaneously, but for clarity, we refer to the agents as target and rater in the context of the interaction being discussed.

We apply our desiderata to a diverse group of trust and reputation mechanisms from the literature. Our desiderata require a utility model, so we have chosen reputation systems that either explicitly define agents' utilities or can be augmented with a utility without further significant assumptions. In each case, we pair off a rational target against a rater as defined by the specific trust or reputation mechanism. We primarily focus on the interaction between two agents, but we examine a few larger settings.

We find that the desirable and undesirable behaviors vary across the mechanisms, validating that our desiderata are granular enough to distinguish differences between models. The general mechanism proposed by both Hazard [2008] and Smith and desJardins [2009] exhibits the most favorable results of those studied when faced with pure moral hazard, although this mechanism does not adapt to a continuous range of behaviors as easily as some other systems do. Moral hazard was more prominently considered in the design of this reputation system when compared to the others we examined, so it is not surprising that this reputation system performs best with respect to our desiderata in a moral hazard situation.

We make an additional contribution in this paper. In order to treat each model as a black box, we present a common conceptual interface for reputation systems. This interface consists of two functions reflecting the two fundamental features of a reputation system.

An update function, used by a rater or central reputation system, which returns a target's new reputation (when participating in the reputation system under consideration) after the target has performed a specified action.

A payoff function, which returns the reward that a target can expect (under the reputation system under consideration) for performing a specified action given its current reputation.

In essence, each reputation system implements the above two functions. In our study we make use of these functions as a programming interface to uniformly incorporate the various reputation systems.

We find that the main limitations of our methods are the computational complexity of finding the optimal strategy for the strategic agent and applying the model to reputation systems that are tightly coupled with complex interaction systems. The results of our desiderata are sensitive to the environment and our desiderata require an explicit utility model for the agents.

The remainder of this paper is organized as follows. First, we discuss the related work in comparing reputation systems in Section 2. We formalize the agent interactions and generic properties of a reputation system in Sections 3 through 3.1 and then discuss how we apply of dynamic systems theory in Section 3.2. In Section 3.3, we discuss agent behavior, focusing on how we employ rational agents. Next, we formalize our method of applying our desiderata (Section 3.4). We describe a basic interaction model with moral hazard in Section 4.1 that we use to evaluate reputation systems. After describing our selection of models from the related literature for comparison in Section 4.2, we demonstrate how our methods and desiderata

are applied to the basic Beta model in Section 4.3. We then apply our desiderata to the chosen models from the related literature in Section 4.4. Next, in Section 5, we discuss the strengths and limitations of our desiderata, along with details of how they may be applied to different systems. Finally, we draw conclusions and discuss future work in Section 6.

2 Current Methods of Evaluating Reputation Systems

The ART testbed [Fullam et al., 2005] is a domain-specific problem for the domain of art purchases designed to test reputation systems. ART is useful for comparing reputation systems in a situated environment. However, the ART testbed suffers from some limitations as a general purpose test to compare reputation systems. One limitation is that the ART testbed does not always align incentives between obtaining a good reputation and increased utility [Sen et al., 2006]. The ART testbed also suffers from issues of ambiguity in agent valuations and capabilities, and being limited to a small number of agents [Krupa et al., 2009]. The domain-specific models in ART are both a strength and a limitation. The strength is that ART adds a practical realism to the measure, but the limitation is that the results depend not just on agents' reputation models, but also on how agents model their interactions and the environment outside of reputation. Our methods are domain independent, isolating the dynamics of the reputation system.

Altman and Tennenholtz [2008] take an axiomatic approach to ranking systems. They prove that, in a multiagent system in the context of aggregate ratings, independence of irrelevant alternatives is mutually exclusive with transitivity. An axiomatic system can yield strong proofs, but realistic models or models with complex interactions often preclude strong results with such modeling due to intractability. Our desiderata treat a reputation system as a black box, which extends its applicability into the realm of reputation systems that use complex computations tailored to specific requirements.

Sybil attacks, that is, agents creating pseudonyms in order to artificially manipulate their or others' reputation, are a frequently studied attack on reputation systems. Resnick and Sami [2007; 2008] use an information theoretic approach to derive worst case bounds on the damage an agent can wreak. Their method limits the amount of influence an agent can wield, but does not account for temporally strategic agents, and focuses on Sybil attacks using randomized actions. Conversely, our desiderata focus on temporally strategic agents.

Besides the aforementioned exceptions, the related literature on reputation systems typically compares a performance measure, often utility, of agents under a specific set of defined attacks for each reputation system. Two surveys indicate the widespread use of this technique. Jøsang et al. [2007] enumerate attacks and other problems, as well as corresponding solutions in the literature. Hoffman et al. [2009] compare reputation systems by which particular attacks their systems address.

Of the attacks employed in the related literature, the most common are agents that behave badly a random percentage of the time [Kamvar et al., 2003; Huynh et al., 2006]; build up a reputation by behaving positively and then "spend" it by behaving badly [Srivatsa et al., 2005; Kerr and Cohen, 2009; Salehi-Abari and White, 2009]; open new accounts to reset reputation [Kerr and Cohen, 2009]; launch Sybil attacks [Kerr and Cohen, 2009; Kamvar et al., 2003; Sonnek and Weissman, 2005]; collude with other agents [Kamvar et al., 2003; Sonnek and Weissman, 2005; Srivatsa et al., 2005]; and change behavior based on the value of the transaction [Kerr and Cohen, 2009]. In contrast, instead of devising attacks solely by intuition, we examine the entire strategy space.

Some of the related work empirically compares more than one reputation system, but such studies comprise a small minority of the related work. Kerr and Cohen [2009] and Sonnek and Weissman [2005] compare several systems across a wide range of attacks, having developed reputation systems to address the

weakness of others. Of the remaining literature that empirically compares reputation systems, many papers compare three or fewer other systems [Huynh et al., 2006; Salehi-Abari and White, 2009]. We postulate that this is in part due to interoperability difficulties between the reputation systems and how some papers do not adequately specify the relationship between valuations, performance, and reputation, thus requiring major assumptions about each reputation system. We address this challenge by presenting a common conceptual interface for reputation systems and discussing how some reputation systems may be implemented using the interface.

General prescriptive desiderata have also been explored in related work [Dingledine et al., 2000; Huynh et al., 2006; Kamvar et al., 2003; Zacharia and Maes, 2000] to guide interaction design and compare reputation systems. Desiderata for trust and reputation systems are not as straightforward [Dingledine et al., 2000] because trust and reputation are supplemental to *primary interaction mechanisms*. A primary interaction mechanism is one, such as a market, that affects agents’ utilities directly. In order for reputation to work, agents must be long lived, ratings must be captured and distributed, and ratings from the past must guide future decisions Resnick et al. [2000].

3 Reputation Dynamics

We represent the attributes of an agent, that is, its *type* including utility functions, valuations, abilities, and discount factors, as $\theta \in \Theta$. An agent may know its type and may keep aspects of their type as private information. The set of all possible agent types, Θ , is dependent upon the system under study. We make no specific assumptions about the space of Θ and simply use θ as a parameter, treating the internals of the reputation system as a black box.

The main purpose of a reputation system is to increase the accuracy of beliefs each agent has about each other agent’s type. An agent’s reputation is a public projection of θ , i.e., it reflects the beliefs of other agents about it. This paper focuses on how an individual rater would assess a given target, and how that rating would affect the target’s ability to gain utility in the future. We use the term *reputation* because in our analysis it provides the elements of what would be the target’s reputation. We denote an individual rater’s belief of a target’s type generically as $r \in R$. We emphasize that whereas r may include information aggregated from the system or other agents, r is the reputation of a target as viewed by a single rater. The domain of r , R , is defined by the reputation system under examination. The domain may be as simple as a nonnegative scalar or as complex as the complete set of possible interaction histories with all details. For the formalisms in this paper, we assume R to be a normed metric space Goffman and Pedrick [1983]. However, all of the metrics and results may be applied using their discrete counterparts. We use the discrete methods when evaluating some existing reputation models.

A target’s reputation is computed by measuring outcomes of direct interactions and by obtaining and aggregating other raters’ experiences and beliefs. The manner by which a rater updates its ratings of a target drives the dynamics of the reputation system. If a rater a rates a target b as r_t at time t , then after a and b interact at time $t + 1$ (or a learns something about b from another rater), a will rate b as r_{t+1} . For example, suppose a currently believes b ’s reputation to be r_t , that b sells high-quality products. If a purchases a product from b at time $t + 1$ that turns out to be of low-quality, a updates its belief of b ’s reputation to r_{t+1} , that b sells low-quality products. Here $r_{t+1} < r_t$. We use r' to indicate the rating after an action or transmission of information has occurred, which is synonymous with r_{t+1} in the case of discrete time.

3.1 Constructing the “Next Reputation” Function, Ω

The idea of this paper is to evaluate reputation systems using a consistent methodology as follows. Given a reputation system, first determine the Ω function that maps an agent’s current reputation to its next reputation. Once Ω is defined, evaluate properties of Ω to understand key properties of the reputation system, especially with regard to its dynamism and convergence when faced with a rational target.

The target chooses how to behave given the environment, its own type, and the specific reputation system employed. The idea is that the target would behave a certain way, taking its current reputation into account when evaluating its decision. This behavior would cause the rater to assess the target a certain way. Based on the specific reputation system, the rater would adjust the reputation of the target appropriately after an observation or new information. Hence the target’s reputation would be mapped from its pre-action value, r , to its post-action value, r' , based on the target’s type, θ , the parameters of the interaction, $g \in G$, the environment, $\psi \in \Psi$, and the reputation system, $\xi \in \Xi$. To capture the above intuitions, we define the function $\Omega : \Theta \times G \times \Psi \times \Xi \times R \mapsto R$ that represents how the reputation of a target changes after an interaction. The target’s decision process is fully captured within the inputs to Ω .

To enable uniformity in assessing different reputation systems, we assume that the rater is rational and patient, and follows the (typically nonstrategic) actions as prescribed by the reputation system under examination. This means that a rater does not lie about reputations unless it is part of the process of the reputation system being examined. For simplicity, we consider the rater’s utility function as a parameter of the interaction. The rater’s utility function is largely governed by the payoff function, which is an input to our desiderata, either as prescribed by the reputation system, as modeled from the interaction environment, or as is used by the actual raters in the system. This is clearly an idealization because in most settings the raters are not strategic agents. However, the idealization systems yields baseline measures of quality and enables us to compare reputation systems.

When making an observation, a rater may also pass information to other agents, either directly or through a centralized mechanism. An agent’s reputation can change with respect to a given rater without a direct interaction. Other than evaluation with a couple of reputation systems in Section 4.4.1, we focus on interactions between two agents. Therefore, for clarity and brevity, we do not explicitly model asynchronous agent communication in our formalisms. We leave this to be handled by the target’s utility function as a change to the environment or as collapsed into an update to the target’s reputation with respect to other agents.

Because a target’s type includes the target’s utility function and decision model, the target’s action can be computed from its type and the other parameters to Ω . Therefore, Ω does not require a parameter for the target’s action.

A target’s decision model must include all actions available to the target. The actions depend on the interaction model employed to evaluate the reputation system. Examples of actions are whether to pay another agent, what quality of item to produce; whether to close the current account and open a new one to reset the agent’s reputation; whether to lie when rating another agent; and whether to open pseudonymous accounts controlled by the target itself to manipulate its own reputation (known as Sybil attacks).

Deciding which parts of a reputation system belong in the Ω function and which parts belong in its parameters is fairly straightforward. Anything that is agent-specific, such as valuations, capabilities, and discount factors should be an attribute of θ . Anything that is common or fixed across all agents, including the processes that define costs and interactions, can be incorporated in the environment, ψ . Attributes which may change from one interaction to another should be specified in g , and the attributes’ domains should be specified by the environment. The mechanisms of the reputation system itself should be incorporated into Ξ and into the Ω function itself.

Throughout this paper, we focus on the process of matching agent types to reputations and how an agent can strategically manipulate a reputation system. When evaluating reputation systems and describing our desiderata, we hold the environment, interaction, and reputation system constant. As the other parameters are held constant, we assume all else remains equal across these interactions, such as the agent relationship topology, valuations, payoffs, game parameters, and probabilities. Our desiderata treat Ω as a black box. For brevity and clarity, we therefore omit the parameters held constant and write a target’s reputation update after the target makes a decision as $r' = \Omega_\theta(r)$.

We make no assumptions about how or whether an agent’s type can change over time. However, when evaluating a reputation system, we hold agents’ types constant; the results from the desiderata indicate how well the system would adapt to changing agent types. Whenever an agent changes its type, a reputation system that meets the desiderata of ACCURACY and CONVERGENCE would catch up with such an agent quickly.

3.2 Fixed Points and Reputation Functions

Because reputation systems are supposed to accurately measure targets’ reputations, a desirable reputation system should yield stable reputations when the targets themselves remain stable. For example, a desirable reputation system should recognize a seller that provides a good product at a low price with a good reputation. Conversely, an undesirable reputation system would be one where a good seller might receive a good or bad reputation only because of luck or strategic reputation manipulation by other agents. An agent’s reputation should follow its type, meaning that a stable agent’s reputation should arrive at a fixed point, ideally corresponding to its type.

A fixed point of a function is where the output of the function is equal to the input. Fixed points are a cornerstone of dynamical systems theory [Devaney, 1992]. The properties of fixed points, such as whether and how they attract or repel, govern the dynamics of systems that have feedback. A reputation is a fixed point if $r = \Omega(r)$, which means that if the reputation were to take the exact value of r , the target’s reputation would remain at the same value after subsequent interactions in an unchanging environment.

The set of fixed points of Ω_θ is $\{r \in R : r = \Omega_\theta(r)\}$. We define the function χ , which yields the stable fixed point, if one exists, of a reputation system for a target of type θ , as

$$\chi(\theta) = \lim_{n \rightarrow \infty} \Omega_\theta^n(r_{\text{initial}}), \quad (1)$$

where Ω_θ^n means that the function Ω_θ is iterated n times. $\chi(\theta)$ depends on r_{initial} , which is the a priori belief that a rater has of a target, given that the rater has no information about the target other than the fact that the target exists. The r_{initial} value is explicitly defined in some systems, and in others it can be assumed to be the expected value over the probability distribution of possible reputations. For example, Sporas defines r_{initial} to be 0, the worst reputation in the domain of $r \in [0, 3000]$ [Zacharia and Maes, 2000]. However, the raters may have differing a priori beliefs or have misinformation about the targets, leading to differing initial reputations. The desiderata of CONVERGENCE and UNAMBIGUITY, described in Section 3.4, address these challenges by saying that a reputation system should ideally have only one fixed point and the reputation should converge toward that fixed point.

In some reputation systems, the limit expressed by $\chi(\theta)$ may not exist. This can be caused by a lack of fixed points, particularly if the domain of possible reputations includes reputations which are impossible to attain or if Ω_θ contains discontinuities with large gaps. The limit expressed by $\chi(\theta)$ may also not exist if the reputation system has a repelling (unstable) fixed point and the reputation never converges to single value. When a target’s reputation oscillates around a single value (i.e., Lyapunov stable with a periodic, toroidal,

or chaotic orbit), we can use that fixed point as the value for $\chi(\theta)$ to apply our other desiderata, noting the caveat that an agent’s reputation will never reach the fixed point, only approximate it. A reputation system could conceivably have multiple fixed points around which a strategic target’s reputation will orbit. The appropriate value for $\chi(\theta)$ in this case is unclear and a marked weakness of the reputation system, but we have not encountered this behavior in any of the reputation systems we examined. We further examine repelling fixed points when discussing the CONVERGENCE desideratum in Section 3.4.

Noise in the environment or stochastic agent strategies can also prevent a reputation system from converging to a fixed point. However, given enough Monte Carlo simulations and analysis, the expected values, moments, and statistical significance can all be propagated through our framework and desiderata. Rather than finding a fixed point, the result will be a stationary stochastic process.

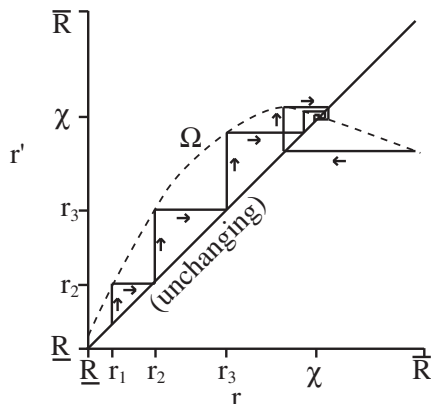


Figure 1: Dynamics of a reputation system.

Figure 1 shows an example “cobweb” diagram as used in dynamical systems theory [Devaney, 1992] for a reputation system with $R = [\underline{R}, \bar{R}] \in \mathfrak{R}$. Because we apply cobweb diagrams to reputation systems throughout this paper as a basis for discussion, we now briefly describe how to read such diagrams. For simplicity in graphical illustration of concepts, we focus on real scalar reputations and real scalar projections of nonscalar reputations throughout this paper, with \underline{R} representing the worst possible reputation and \bar{R} representing the best possible reputation. The bounds of possible reputation values depend on the reputation system and need not be finite. For our discussions of reputation systems with real scalar values, an unbounded maximum reputation means $\bar{R} = \infty$.

In our application of cobweb diagrams, the horizontal axis represents the target’s current reputation over the domain of possible reputation values. The vertical axis represents r' , with the dashed line representing the target’s next reputation after performing the action as governed by its type, $\Omega_\theta(r)$. The diagonal line represents unchanging reputation and helps identify fixed points. A fixed point exists wherever an Ω_θ function intersects the diagonal line.

Figure 1 shows two starting points to illustrate how the reputation changes over time. Suppose a target has a bad reputation, as indicated in this illustration as a low value where the stair-step line starts on the bottom left. What constitutes a bad reputation depends on the specific reputation system (and the associated decision model of the targets), but generally we say a target has a bad reputation if another rater believes the target will likely offer poor-quality products or otherwise behave in an undesirable fashion (we return to this point in Section 3.3). The target’s subsequent reputation, that is, the target’s reputation after performing its next action, is the value on the dashed line above the horizontal position indicating the target’s current

reputation. This value is then used as input for the next interaction. The target begins with reputation r_1 and its strategy leads it to perform actions that lead its next reputation to be calculated as r_2 —and so on, through the series of steps in the diagram. We can find each successive reputation by moving horizontally to the diagonal line and then moving vertically to the new location on the dashed line. In this example, the reputation converges to the (only) fixed point marked by χ on each axis. If the target’s reputation somehow becomes higher than the fixed point in this graph, the strategic target would “expend” a small amount of its reputation, for example, by providing poor service. As a result, the target’s reputation would be lowered to lie below the fixed point. However, once the reputation is below the fixed point, the target would behave nicely and continue to rebuild its reputation back up to the fixed point. Then it would expend it again, and so on.

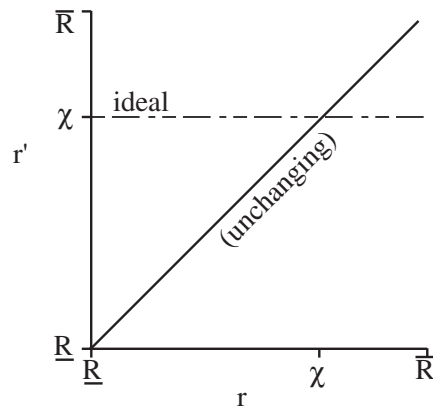


Figure 2: An ideal reputation system.

Figure 2 depicts an ideal reputation system. The horizontal line represents the ideal case as expressed by $\Omega_\theta(r) = \chi$. This represents an ideal reputation measurement system because the reputation is measured accurately in one shot regardless of what the target’s previous reputation was. This ideal case is only useful if χ depends appropriately on θ —in other words, if χ accurately reflects the type of the target. A reputation system that always returns the same reputation regardless of behavior may be perfectly precise, but it would be neither accurate nor useful.

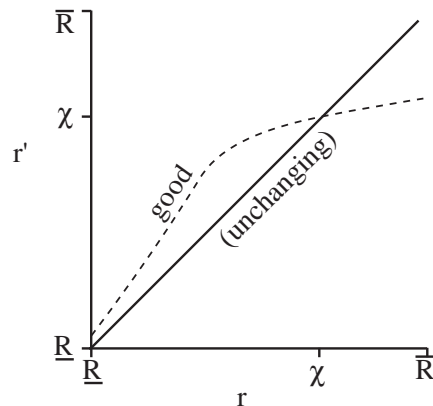


Figure 3: A good reputation system.

The dashed line labeled “good” in Figure 3 represents a reputation system that converges to a fixed value regardless of other raters’ previous beliefs. Targets whose reputations are greatly undervalued build their reputations slowly, whereas targets with overly inflated reputations slowly converge on an appropriate reputation without their reputation values bouncing around. These dynamics may be observed by using the same stepping method as described for Figure 1. Such a reputation system might be indicative of an e-commerce setting where targets with poor reputations charge low prices for decent-quality products and, as they build up their reputation, they begin to charge more for their offered level of quality. In this reputation system, if the target’s reputation is overinflated, it may take advantage of the situation by possibly lowering the quality of its product slightly or raising the price, until it achieves its equilibrium fixed point reputation.

3.3 Agent Behavior

The key concepts in this paper, particularly the desiderata introduced in Section 3.4, directly apply to any type of agent decision model. One example of a decision model is an agent that plays strategies based on a stochastic process. Another example is a malicious agent whose utility function increases with the utility loss of another agent. However, we primarily focus on rational agents.

When moral hazards exist in an interaction setting, *strategic agents* can be a major threat to a reputation system. A strategic agent will do whatever actions lead toward achieving a goal, and would thus exploit any mechanism or manipulate its reputation if doing so helps achieve the goal. A rational agent is a type of strategic agent that evaluates all possible future actions and payoffs, which often must be approximated due to uncertainty and computational complexity, then chooses the immediate action that will lead it to the largest total payoff (we discuss details of this for our particular experimental evaluation in Section 4.1). Although the resilience of a reputation system against strategic agents indicates how well the reputation system may fare in an open real-world setting, much of the related literature on reputation systems does not discuss strategic agents. Of the papers that do discuss strategic agents (e.g., [Kamvar et al., 2003]), only a minority formally model strategic agents (e.g., [Jurca and Faltings, 2007]). A rational agent may maximize its expected utility over its expected lifespan or use intertemporal discounting. Thus a rational agent’s Ω function is the path of reputation that maximizes its utility.

To consistently quantify the comparison of reputation values in relation to agent types, we focus on the case when a target is faced by an *ideally patient strategic (IPS)* agent. We define an IPS agent as a rational agent that is indifferent to the time of when a specific utility change will occur.

Our motivation for considering IPS agents is as follows. Since the idea of reputation is to help select agents for future interaction based on their expected future behavior, it is natural that we rate targets in a manner that places substantial weight on the future utility of the rater. Specifically, if agent a is interacting with an impatient agent b , then a may perform actions (that affect b) that would be considered socially detrimental to a patient agent. For example, consider two agents in a situation where they can gain utility only by cooperating and offering each other favors. Agent a might not provide a favor to b if a believes that b is not patient enough to return a favor to sustain a mutually beneficial long-term relationship [Hazard, 2008]. In this case, any observing agents (or centralized rating mechanism) should not necessarily observe a as being impatient or having a low reputation because a is simply protecting itself against agent b . If b were measuring the reputation of a as the target, b would be unable to distinguish between a myopically greedy agent and an agent that was simply protecting itself against b ’s behaviors. By measuring a target against an IPS agent, we can ensure the target does not need to apply any “self-defense” measures because the target has perfect knowledge that the IPS agent will not attempt to take advantage of it for short-term gain. Further, by definition, an IPS agent values a longer running good reputation more than a less patient agent.

A patient agent is also more useful for comparing reputation values than an impatient agent because a patient agent generally offers a larger possible range of behaviors. This notion is supported by the economics literature (e.g., Fudenberg and Levine [1992]). Suppose b is a reference agent, a rater by which we are measuring a property of target a . If b is impatient, then b would attempt to take advantage of a whenever doing so offered a large immediate payoff, regardless of a 's type and behavior. Conversely, if b is ideally patient, then b 's behavior will reflect b 's belief of a 's type, providing a measurement of a 's type.

Suppose rater b is rational and is interacting with a target a that has type θ_a . Rater b maximizes its total utility, U_b , by controlling its strategy, σ_b , which is a set containing a specific action at each time t , $\sigma_{b,t}$. At each time step, b receives utility $v(\theta_a, \sigma_{b,t})$, which is a function of b 's strategy, b 's type, and a 's type, from which a 's optimal strategy may be derived. For an IPS agent, the function v should be chosen to represent typical agents in the system, that is, to represent average valuations and capabilities, or be endowed with capabilities and valuations the designer feels represent a good benchmark for the system. In this paper, we use the same valuations across all agents, including IPS agents.

Definition 1 We define an ideally patient strategic agent (IPS agent), b , as having an infinite time horizon such that b maximizes its average expected total utility as a function of any agent a 's type, θ_a , as the time horizon goes to infinity as

$$E(\bar{U}_b(\theta_a)) = \max_{\sigma_b} \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \sum_{t=0}^{\tau} v_{\theta_b}(\sigma_{b,t}, \theta_a) = \max_{\sigma_b} \lim_{\gamma \rightarrow 1} (1 - \gamma) \sum_{t=0}^{\infty} \gamma^t v_{\theta_b}(\sigma_{b,t}, \theta_a). \quad (2)$$

We use an IPS agent's utility function, given an environment, interaction, and so on for ordering agent types by preference. If an IPS agent b prefers target a to target c , that means $E(\bar{U}_b(\theta_a)) > E(\bar{U}_b(\theta_c))$. In the case of simple favor transactions with pure moral hazard, the IPS agent prefers targets with higher discount factors because such targets may yield higher payoffs. For example, the IPS agent may achieve higher payoffs with a patient agent via a trigger strategy, where both agents would follow some schedule of actions and be punished for deviation [Axelrod, 2000], because a patient agent would be willing to sacrifice short-term loss to achieve the long-term gain from the schedule of actions. When agents offer products of differing quality for differing prices with pure adverse selection, the IPS agent prefers agents whose products maximize value over time.

Evaluating the average expected total utility of an IPS agent is not necessarily always an easy task. Numerical evaluation methods are useful for approximating the limits. Because the process of backward induction generally does not apply to infinite horizon games, finding the expected utility as $\gamma \rightarrow 1$ is a viable approximation as long as the set of interactions is small enough that searching through enough plies of interactions is tractable.

3.4 Reputation System Desiderata

Reputation systems may be useful and effective even if their behaviors are not close to ideal. This section examines what makes one reputation system more desirable than another and what can render a reputation system ineffective.

Consider the line labeled *good* in Figure 3. The strategic target would eventually attain its fixed point reputation. However, if Ω_θ yields similar curves for all θ , a rater would not be able to distinguish among different targets based on variations in their reputation because they would all end up with the same reputation value. This may be acceptable when the target has an extremely favorable type, but if other targets' types yield the same structure, then a strategic target may be able to gain a better reputation than it deserves. This

is not to say that a system in which all targets achieve a good reputation is necessarily bad. A mechanism that incentivizes targets to always behave in a socially beneficial manner, regardless of their type, can be desirable. However, if target a has a better reputation than target b , then a trustworthy agent c should expect a to behave at least as well as b in interactions, all else equal, regarding c 's own utility. Relating this concept back to the ideal reputation system in Figure 2, the horizontal line representing target a 's type would be at a more desirable reputation value than that of target b 's type.

One reputation is better than another if, with all else equal, the rater expects greater utility interacting with a target with the better reputation. For an IPS agent, c , entering a relationship of repeated interaction with agent a , this utility is a function of the other agent's type, θ_a , $E(\bar{U}(\theta_a))$. A regular rater, however, would not know a 's type, but only its reputation, and would only evaluate a single transaction. We write a rater b 's utility of entering an interaction with a as $u(\chi(\theta_a))$. The function u is the payoff function that yields the value of a single transaction for a given reputation, which is a property of the reputation system under examination.

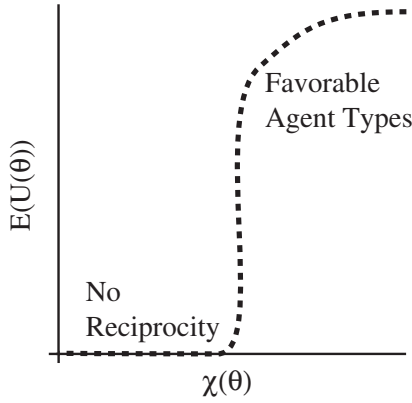


Figure 4: Parametric plot of $E(\bar{U}(\theta))$ and $u(\chi(\theta))$ with respect to θ .

Figure 4 shows how an IPS agent's utility changes with respect to the fixed point reputations of a one-dimensional agent type. In this example, the IPS agent would not interact with unfavorable agent types because they would try to reduce the IPS agent's utility for their own gain. At some threshold of θ , an agent would enter a mutually beneficial relationship with an IPS agent, with more favorable agents bringing greater utility to the IPS agent. If this parametric plot were not monotonic, an agent with a high reputation would have a lower expected utility to an IPS agent than an agent with a lower reputation.

Desideratum 1 MONOTONICITY: *If, to an IPS rater c , target a 's type is preferable to target b 's type, then a 's asymptotic reputation should be greater than b 's reputation. More formally, a reputation system is **monotonic** if $\forall \theta_a, \theta_b \in \Theta : E(\bar{U}_c(\theta_a)) \geq E(\bar{U}_c(\theta_b)) \Rightarrow u(\chi(\theta_a)) \geq u(\chi(\theta_b))$. However, if c is indifferent across all agent types, that is, $\forall \theta_a, \theta_b \in \Theta : E(\bar{U}_c(\theta_a)) = E(\bar{U}_c(\theta_b))$, then the reputation system is **nondiscriminatory**, a generally undesirable subset of the otherwise desirable monotonic property.*

As in Figure 2, an ideal reputation system would enable a rater to assess a completely unknown target's reputation perfectly after one interaction. The closer Ω is to $r' = \chi$, a horizontal line for one-dimensional reputation measures, the lower the error is between the target's current reputation and what it asymptotically approaches. We define this error as on the domain of possible reputations, R , as follows.

Definition 2 We define reputation measurement error, $\epsilon \in [0, 1]$, at some reputation r for a target of type θ as the distance between a new reputation $\Omega_\theta(r)$ and the asymptotic reputation χ , normalized with respect to the maximum distance between any two reputations, as

$$\epsilon_\theta(r) = \frac{|\chi(\theta) - \Omega_\theta(r)|}{\max_{x,y \in R} |\Omega_\theta(x) - \Omega_\theta(y)|}. \quad (3)$$

Definition 3 We define average reputation measurement error (ARME), $E(\epsilon_\theta) \in [0, 1]$, as the expected value of reputation measurement error for target type θ across all possible beliefs of reputation, normalized over all possible reputations, R , with the exterior derivative of R , dr , as

$$E(\epsilon_\theta) = \frac{1}{\int_R dr} \int_R \epsilon_\theta(r) dr. \quad (4)$$

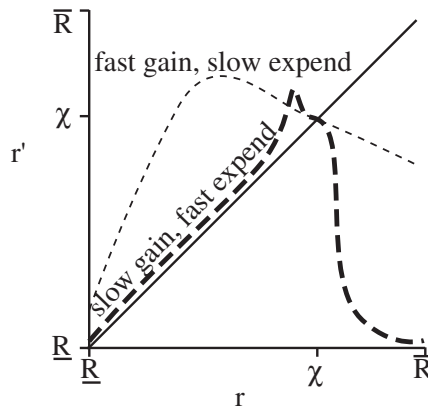


Figure 5: Reputation systems with different amounts of error.

Figure 5 shows two reputation systems, each with one fixed point and the same derivative at the fixed point. In the reputation system shown by the line labeled *fast gain, slow expend*, targets with low reputations quickly improve their reputation, but the reputation can overshoot and would oscillate as it approaches χ . A reputation system producing the line labeled *slow gain, fast expend* would have targets gain reputation more slowly than *fast gain, slow expend*, and targets that gain overly valued reputations would quickly expend a significant amount of reputation; some targets would cause large oscillations in their reputation, possibly for a significant period of time before their reputation stabilizes, if ever. An example of *slow gain, fast expend* is the recent major Ponzi scheme by Bernard Madoff, where he had gained a strong reputation throughout his career and allegedly used his reputation to build the Ponzi scheme.¹ Qualitatively, the *fast gain, slow expend* reputation system is generally preferable to *slow gain, fast expend* because it is more stable and accurate. The ARME provides a quantitative comparison, yielding a lower error for the *fast gain, slow expend* reputation system.

Although ARME gives the error for a single target type, an important purpose behind a reputation system is to deal with different target types. One reputation system may yield low error with targets of bad reputations whereas another reputation system may yield low error with targets of good reputation. Further, a system may have mostly good or mostly bad agents, so a reputation system designer should evaluate and compare reputation systems based on the expected mix of target types.

¹<http://www.sec.gov/news/press/2008/2008-293.htm>

Desideratum 2 ACCURACY: The average reputation measurement error, $E(\epsilon)$, should be minimized with respect to the believed distribution of target types, represented by the probability density function $f(\theta)$, where $E(\epsilon) = \int_{\Theta} f(\theta) \cdot E(\epsilon_{\theta}) d\theta$.

Whereas ARME gives an indication as to how the reputation system performs across all reputations, it does not give an indication as to how the system performs when a rater's belief of another's reputation is somewhat accurate. To address this situation, we now discuss reputation dynamics around a fixed point.

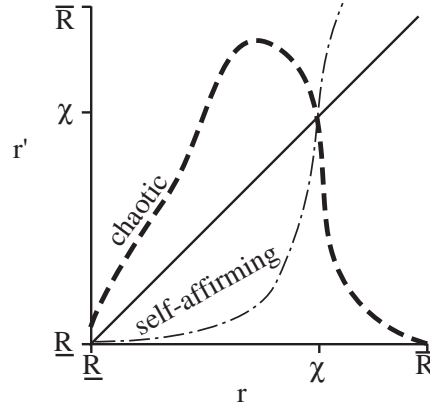


Figure 6: Reputation systems with large derivative magnitudes at the fixed point.

A fixed point is said to be *attracting* if the dynamical system asymptotically converges to the fixed point when starting near enough to it. A fixed point may also be *repelling*, meaning that the dynamical system diverges from the fixed point unless exactly at the fixed point. An example of a repelling fixed point is the fixed point of the line labeled *self-affirming* in Figure 6. If a reputation system has a single fixed point, then over time the accuracy of a target's reputation increases for an attracting fixed point and decreases for a repelling fixed point. Dynamical systems may also be attracted to or repelled from a periodic cycle of a number of points, or end up chaotic, meaning that the value jumps around within a region in an unpredictable manner [Devaney, 1992].

A fixed point can be attracting on one side and repelling on the other if Ω is tangential to the line $r' = r$ or if the derivative is not continuous at the fixed point. Systems whose derivatives are not continuous at their fixed points can act similar to systems with no fixed points because the reputation moves in only one direction in each case. However, if a target's reputation asymptotically approaches the fixed point from the attracting side but does not cross the boundary of the fixed point, then the system can still exhibit stable reputations.

In Figure 6 the line labeled *self-affirming* depicts a reputation system in which a rater's eventual belief of a target is completely dependent on its initial belief due to the repelling fixed point. By tracing the feedback of this function, a reputation below χ would eventually end up at \underline{R} and a reputation above χ would eventually end up at \bar{R} . Such a mechanism is not generally useful for measuring reputation, but may nevertheless be useful as an interaction mechanism if

- prior beliefs begin at specified values, e.g., when all agents participating in an online auction automatically start with a neutral reputation;
- better reputations incentivize targets to perform in a more socially beneficial manner, e.g., an online auction that explicitly awards higher payoffs to agents with better reputations; or

- it is otherwise effective in alleviating moral hazard, e.g., a system in which agents with a low reputation are permanently banned.

The curve labeled *chaotic* in Figure 6 shows a repelling fixed point that causes a target’s reputation to remain persistently unstable. Below the fixed point, the target’s reputation grows quickly. Once the target’s reputation is above the fixed point, the target’s best strategy is to take actions that quickly reduce its reputation. A reputation system exhibiting this behavior would likely be ineffective because a target’s current reputation is usually meaningless with regard to its type. An example of such a system is a peer-to-peer file sharing service where agents first must upload content before they can download content. In this case, the agent must first build up its reputation by uploading files to other peers, and then the agent can expend its reputation by downloading. An agent’s reputation, that is, the amount of data uploaded or downloaded, functions similar to a currency.

Whether a fixed point is attracting or repelling depends on the derivative at the fixed point [Devaney, 1992]. A fixed point is an attractor if Ω is a local contraction mapping if its Lipschitz constant, the minimum bound of the scaling factor between successive iterations, is less than 1. As we are looking at local dynamics, we can express this constraint on the Lipschitz constant as the maximum component of the gradient as $\|\nabla\Omega(r)\|_\infty < 1$ at χ , where a target’s reputation eventually converges provided no other fixed points exist that change the dynamics.² If $\|\nabla\Omega(r)\|_\infty > 1$ at χ , then the fixed point repels. When multiple fixed points exist, repelling fixed points can create periodic or chaotic dynamics. If $\|\nabla\Omega(r)\|_\infty \approx 1$ at χ , then the reputations do not change on the fixed point. In this case, a target is incentivized to perform at whatever reputation level it happens to be at, that is, at whatever level the rater believes it to be at; the target’s performance at that level simply reinforces the rater’s belief about it.

Attracting fixed points need not converge in a stable manner; a negative derivative causes a reputation to oscillate about the fixed point whereas a positive derivative approaches the fixed point from one side. The closer to zero the derivative is, the faster the reputation approaches the fixed point and the quicker the reputation gains accuracy.

Desideratum 3 CONVERGENCE: *At the fixed point, $\chi(\theta)$, the sequence of utility maximizing reputation values must be attracting and should converge quickly, that is, $\|\nabla\Omega(r)\|_\infty|_{r=\chi(\theta)}$ must be less than 1 and should be minimized.*

Although any number of fixed points may exist for a given target type in a given reputation system, the ideal number is one. If zero fixed points exist, then the reputation values themselves are asymptotically meaningless. In order for a system to have no fixed points, one of a couple of specific situations must occur. One is if the reputation is unbounded such that a target can attain an arbitrarily high reputation and the reputation remains high even if the target behaves in a manner that should yield a low reputation. An example of a reputation system yielding this behavior would be one where only positive encounters were recorded; because negative encounters are ignored, a target could simply provide enough positive experiences to build a good reputation and provide many negative experiences to boost its own profit. Another case where fixed points may not exist is when Ω is discontinuous, such as a reputation system where agents are incentivized to oscillate between very good and very bad reputations.

When targets’ reputations are unbounded and the mechanism has no fixed points, all targets could end up with an unboundedly growing reputation, as in the aforementioned case represented by the *saturating* line in Figure 7. If, for all target types, Ω is completely below the diagonal except for the lowest reputation

²For a scalar reputation, this can be expressed more simply as $|\frac{d\Omega}{dr}| > 1$.

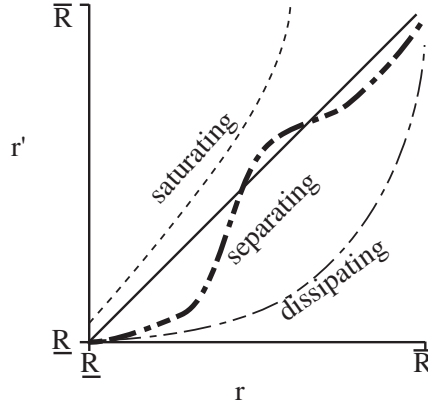


Figure 7: Reputation systems without meaningful fixed points.

value, as shown by the line labeled *dissipating* in Figure 7, then all targets would eventually end up with the worst possible reputation. The dissipating case is similar to the saturating case except that a target's reputation continually decreases. Each target's optimal strategy is to always reduce its reputation, leaving the reputation system meaningless outside of a target's a priori reputation. Because a target's reputation and thus payoff are both guaranteed to continually decrease, a reputation system with such dynamics is generally a poor choice from the standpoint of mechanism design. One real-life example of such a situation is certain vendors at tourist traps. If they provide low-quality goods, tourists do not buy from them, so they increase the sales pressure. At some level of sales pressure, enough tourists do buy from such vendors just to get the vendors to stop trying to sell to them, further incentivizing the vendors to increase pressure on selling low-quality items.

If multiple fixed points exist, then the fixed point that is asymptotically achieved depends on the rater's initial beliefs, hence the reputation is ambiguous. Consider the line labeled *separating* in Figure 7. If the target's reputation is above the middle of the reputation domain then the target's reputation converges to χ . If the target's reputation starts below the middle, then it continually receives a lower reputation until it reaches the lowest possible value. Note that this depends solely on the other rater's initial belief; if rater *a* believes target *b* has a low reputation, then such a reputation system exaggerates this incorrect yet self-fulfilling belief. This type of graph might be seen in the following reputation situation. Consider a manager at a business who highly values his initial opinions and does not like to be proven wrong. If the manager believes that a new employee will excel, he might give the employee many more opportunities to excel than to an employee who he believes will not excel. Because of the positive reinforcement in this situation, the manager's initial beliefs may become a self-fulfilling prophecy.

Having multiple fixed points is not necessarily a problem for a reputation system. If a target's reputation cannot possibly get to a fixed point, the fixed point is irrelevant. On an online auction site, for example, the reputation dynamics of targets with low reputations do not matter if the site bans a seller's accounts if the seller's reputation drops below a certain threshold. If the reputation system depicted by the aforementioned *separating* line from Figure 7 starts all targets off with the maximum reputation then the targets may not ever reach the lower region because the graph changes shape accordingly with the target's type (separate diagrams could be plotted for all values of θ to demonstrate this, much like Figure 4). An example of a reputation system that can exhibit this kind of behavior is one that values a long positive history significantly more than recent actions. For a desirable target type, the system might have a fixed point at a high reputation and

another at a low reputation. In order for a target with a desirable type to achieve the lowest fixed point, it might need to irrationally expend effort to make outcomes bad enough to diminish its reputation to the point where it is better to put no effort into the quality of its products. If the target is rational, then this fixed point will not be reached unless the target’s actions are at least partially driven by a stochastic process and the target was particularly unlucky, which may happen on occasion within an environment with a large enough number of agents. In a market with significant competition, few sales, and small profit margins on products, a target with a favorable type but low reputation may not find it profitable to expend the effort required to build up to a higher reputation fixed point.

Even if a reputation system has theoretically inaccessible fixed points, in practice it does not necessarily mean that it is impossible for targets to reach this region; errors and unforeseen cases could make it possible. A shipment may be lost by an intermediate party who denies responsibility or a bug in software can cause a rating to be inaccurate with respect to the target being rated. Therefore, it is most desirable for a reputation system to have one fixed point per target type. If exactly one fixed point exists for a given target type, then the fixed point is the target’s reputation. The ideally descriptive case is when the mapping between type and reputation is bijective.

Desideratum 4 UNAMBIGUITY: *A target’s reputation should be asymptotically unambiguous, that is, $\forall \theta \in \Theta : |\{r \in R : r = \Omega_\theta(r)\}| = 1$.*

4 Empirical Results

We now apply our desiderata to some important reputation systems. We investigate the reputation measurement aspects of each system. For each system, we briefly review the reputation measure it embodies, discuss utility considerations, and then directly evaluate the reputation systems on a simplified transaction model exhibiting moral hazard.

4.1 Experimental Method

We evaluate reputation mechanisms using a simple, stylized interaction mechanism for two reasons. First, we use a simple model to keep the problem of evaluating optimal reputation strategies tractable in order to evaluate the reputation mechanism itself; complex markets can require an intractably large number of evaluations [Fullam and Barber, 2006]. The second reason is that complex markets make it more difficult to isolate the effects of a single target’s strategy [Kerr and Cohen, 2009].

In each round of our interaction model, a rational agent begins with a specified reputation. The rational agent begins in the role of target, choosing whether to offer a favor to another agent in the initial role of a rater that is operating using the reputation system being evaluated. If the rational agent offers the favor, it incurs a cost of c to itself and the other agent would receive a benefit of b . These roles are then reversed, where the other agent chooses whether to offer the rational agent a favor with the same payoffs, and the round is concluded. To show that “gains from trade” are usually possible when agents grant favors to one another, we examine these variables in the ranges of $c \in [1, 12]$ and $b \in [10, 30]$.

We evaluate each system with respect to the above desiderata against rational targets across the range of possible discount factors. A discount factor is how an agent places less value on future events than on present events. Discount factors arise from combinations of factors such as the uncertainty of a future event occurring and external methods of compounding utility (e.g., investments). Discount factors are widely employed in decision models across economics and artificial intelligence [Dellarocas, 2005; Ely and Välimäki, 2003;

Hazard, 2008; Jurca and Faltings, 2007; Saha et al., 2003]. We employ the commonly used exponential discounting method. Using this method, each agent multiplies the expected utility of an expected future event by γ^t , where $\gamma \in [0, 1]$ is the discount factor and t is the time of the event relative to the present. Discount factors can directly affect an agent’s optimal behavior and thus reputation.

As we discussed in Section 3.4, an agent’s patience can affect its behavior and ability to observe behavior in others. Discount factors are a quantitative measurement of patience. A greedy target might rapidly expend its reputation, whereas a patient target may build and retain its reputation. When evaluating reputation systems, we investigate behavior across the range of possible discount factors.

In our simulations, the possible strategies of a target are a series of binary decisions. That is, each strategy is a sequence such as $\langle \text{favor, favor, nofavor, } \dots \rangle$. We limit the length of the strategies we consider via STRATEGYDEPTH, a parameter of the simulation. For this reason, we write the set of possible strategies in a regular expression notation as $\{\text{favor, nofavor}\}^{\text{STRATEGYDEPTH}}$. In our simulations, we set STRATEGYDEPTH such that the 95% of the total utility over the infinite horizon is captured with respect to the agent’s discount factor, meaning $\text{STRATEGYDEPTH} = \lceil \log(1 - 0.95) / \log(\gamma) \rceil$.

Algorithm 1 ComputeNextReputation(raterModel, target, targetReputation)

```

1: bestUtility  $\leftarrow -\infty$ 
2: nextReputation  $\leftarrow$  targetReputation
3: strategySpace  $\leftarrow \{\text{favor, nofavor}\}^{\text{STRATEGYDEPTH}}$ 
4: for all  $s \in$  strategySpace do
5:    $(\text{util}, r) \leftarrow$  ComputeUtilityAndReputationFromStrategy(raterModel, target, s, targetReputation)
6:   if util  $>$  bestUtility then
7:     bestUtility  $\leftarrow$  util
8:     nextReputation  $\leftarrow$  r
9:   end if
10: end for
11: return nextReputation

```

To find the optimal strategy for a given discount factor, we compute the utility gained for each possible strategy of the entire tree of the extended form game, as outlined in Algorithm 1. Each time the rational target is given the opportunity to decide whether to offer a favor, both decisions are followed. This algorithm approximates $\Omega_\theta(r)$ to the depth of the game tree as specified by the constant STRATEGYDEPTH. Because of the intertemporal discounting, each successive decision yields less utility, and so the utilities of infinitely long strategies may be approximated when the future expected utility falls sufficiently close to 0 with respect to the payoffs from the target’s actions in the nearer future. The overall computation of this Markov decision process is exponential in the number of decisions followed. A rational target’s future expected utility for a particular reputation, $U(r)$, can be expressed recursively in its Bellman equation form as

$$U(r) = \max_{\sigma} (u(r, \sigma) + \gamma \cdot U(N(r, \sigma))), \quad (5)$$

where σ is the agent’s action, $u(r, \sigma)$ is the utility it expects to get for a given time step, and $N(r, \sigma)$ is the agent’s new reputation after it performs σ . The agent’s action will be that which maximizes utility for the current reputation, r , that is, the outermost σ . Algorithm 2, which is used on line 5 in Algorithm 1, evaluates this expression for the model-specific functions raterModel.GetNextReputation and raterModel.GetExpectedActionPayoff to find the total utility and next reputation of a target that employs a

particular strategy. Lines 5 through 8 of Algorithm 2 compute the cost that the target occurs if its strategy is to offer a favor for the given timestep, and line 14 computes the benefit that the target receives gets from the rater based on the target’s reputation. Between these two payoffs, line 11 updates the target’s next reputation given its current reputation and most recent action.

The functions `raterModel.GetNextReputation` and `raterModel.GetExpectedActionPayoff` express the entire functionality of the reputation system, encompassing the effects of multiple agents if applicable. The first function, `raterModel.GetNextReputation`, returns the target’s next reputation with respect to the rater, updated from its current reputation by the action it performs. The second function, `raterModel.GetExpectedActionPayoff`, returns the expected payoff that the target will receive given its reputation and whatever parameters are used to determine the benefit. In our particular evaluation scenario, the payoff is independent of the target’s action because the rater does not know the target’s action. However, in other situations, `raterModel.GetExpectedActionPayoff` may be a function of some information about the target’s strategy, for example, if the target’s action contains a publicly observable signal such as the fact that a product was shipped via an impartial third party.

Algorithm 2 `ComputeUtilityAndReputationFromStrategy(raterModel, target, targetStrategy, targetReputation)`

```

1: utility ← 0
2: currentRep ← targetReputation
3: for timeStep = 1 to length(targetStrategy) do
4:   //target plays its strategy
5:   if strategy[timeStep] = favor then
6:     //if the target gives a favor on this timeStep, it loses some utility
7:     utility ← utility - target.γtimeStep-1 · FAVORCOST
8:   end if
9:   //rater reacts and plays its strategy according to the model
10:  //for example, if the target gave a favor above, the rater might respond by raising the target’s reputation
11:  currentRep ← raterModel.GetNextReputation(currentRep, strategy[timeStep])
12:  //depending on the target’s updated reputation, the rater would reward it with a FAVORBENEFIT
13:  //the FAVORBENEFIT would add to the target’s utility
14:  utility ← utility + target.γtimeStep-1 · raterModel.GetExpectedActionPayoff(currentRep, FAVOR-
    BENEFIT)
15: end for
16: return (utility, newReputation)

```

It may be possible to analytically solve some of the models for the optimal solution, but others are quite complex. We thus use a brute force analysis because it works across all models. However, because this brute force analysis is costly, we do not explore the region of rational agents with the highest discount factors (above 0.90 for individual agents and above 0.60 for networks of agents). Unless an unforeseen phase change exists in any of the reputation models with discount factors greater than 0.90, we expect our results should be representative of the higher discount factors.

4.2 Choice of Models

Whereas many reputation systems have been proposed and studied [Ramchurn et al., 2004], little work has directly compared their effectiveness in general terms. From the body of literature, we choose systems based on the following criteria.

- The system measures reputation and does not merely aggregate reputations without specifying how reputation is defined for a given context. Trust propagation is an important topic, but as our reputation measures examine the entire system, agents need some method of measuring trust.
- The reputation as measured either explicitly characterizes the agents' utilities or can be used as a basis for making decisions regarding their interactions.
- The implementation is straightforward and well-defined. This means that we identify papers that provide sufficient information to recreate their model. This also means that we sought models that did not require a large number of abstract measurements and parameters and could be applied to simple interactions without requiring a market.
- The set of systems considered is diverse. To demonstrate the generality of our approach, we consider models based on different principles and philosophies.

Whereas we omitted some models due to the above criteria, this omission does not necessarily mean that our measures cannot be applied to them. Kerr and Cohen's Trunits model [2006] requires a market whereby agents need to have some input or control with respect to their goods' prices. Our simple favor experiment would not adequately explore the Trunits reputation space, but a more complex scenario could meet this need, albeit with the requirement of further computational complexity to evaluate the optimal strategies. Similarly, Fullam and Barber's model [2006] is designed for the complex interactions in the ART testbed [Fullam et al., 2005]. Other models, such as that described by Zhang and Cohen [2007], focus on large-scale aggregation. Many of the models focusing on large-scale aggregation resemble or build upon another model that focuses on individual agents; in Zhang and Cohen's case, their model resembles the Beta model. Sierra and Debenham's information theoretic model [2005] explicitly uses preferences rather than utilities, and is geared toward richer interactions where agents have many possible actions.

The dynamics of a reputation system are greatly influenced by the relationship between reputations, capabilities, and utilities. If a good reputation is expensive to build and maintain, but the difference in utility between having a bad versus a good reputation is small, then even trustworthy agents would not have an incentive to build up their reputation. This is analogous to diminishing returns seen by a company when improving the quality of a product that already meets the standards expected in the marketplace. For example, if agent a in a peer-to-peer environment is requesting a file transfer from agent b , agent a may not notice any difference in service if b 's upload bandwidth is slightly greater than a 's download bandwidth versus if b 's upload bandwidth is ten times a 's download bandwidth. For reputation systems (Beta and Sporas) which do not explicitly provide a utility model, we apply a utility model inspired from empirical results on online auctions.

4.3 Applying Desiderata to Existing Systems: Beta Model Example

The basic Beta model reputation system is a good exemplary case to apply our desiderata because the reputation mechanism itself is simple to implement and understand, yet contains a few minor hurdles with respect to applying our desiderata.

The Beta model is a frequently studied and extended reputation measure [Jøsang, 1998; Jøsang and Quattrocchi, 2009; Teacy et al., 2006; Wang and Singh, 2006, 2007], where agents rate each experience with another agent as positive or negative. Using this method, raters quantize interactions into positive and negative experiences and use a beta distribution to indicate the probability distribution that a target will perform positively in the future. Given a number of positive interactions, α , and negative interactions, β , the expected probability that a future interaction will be positive is $\frac{\alpha}{\alpha+\beta}$, the mean value of the beta distribution. Reputation systems using this approach typically assume that agents are not rational and have an intrinsic probability of performing positively or negatively. A target’s reputation is its expected probability of yielding a positive interaction.

The following steps are the process for applying our desiderata measures to a reputation system. In each step, we use the interaction model specified in Section 4.1 with the basic Beta model as an illustrative example.

1. Determine the update function. For the Beta model, the update function is straightforward with respect to our interaction model. A rater rates the target positively if the target offered a favor, or negatively if the target did not. A rating, r , consists of a tuple of two nonnegative integers: the total number of positive interactions, $i_{P,r}$, and the total number of negative interactions, $i_{N,r}$. The update function, n , for the Beta model can be expressed as

$$n(r, \sigma_t) = \langle i_{P,r} + \sigma_t, i_{N,r} + (1 - \sigma_t) \rangle, \quad (6)$$

where σ_t is the strategy of the target at time t containing 1 if it will offer the favor and 0 if it will not.

When computing an agent’s payoff or plotting an agent’s reputation using the Beta model, we use the belief of a positive outcome, b_P , as the scalar value of an agent’s reputation, as defined by Jøsang [1998]. For a given reputation r , b_P can be expressed as

$$b_P(r) = \frac{i_{P,r}}{i_{P,r} + i_{N,r} + 1}. \quad (7)$$

2. Determine the payoff function. Adding utility to the Beta models is relatively straightforward. Because the transactions are quantized as being positive or negative, we assume that each carries a constant utility. As reputation is the probability that interacting with the given agent will generate a positive transaction, the expected utility is simply the probability of each outcome multiplied by the utility of each outcome. From a strategic agent’s perspective, the main difference between interacting with a single agent using the Beta model and a population of communicating agents using the Beta model is the number of observations any given target will have.

The exact relationship between reputation and price can be unclear in some contexts [Resnick et al., 2006], but Melnik and Alm [2003] have found a multiplicative relationship with sublinear and superlinear terms between reputation and price on eBay, a major online marketplace. To explore some reputation systems further, we apply three utility models with respect to reputation. The first is linear, meaning that a perfect reputation yields full utility, a middle reputation yields half utility, and the worst reputation yields no utility. We also investigate a sublinear relationship, where the scalar representation of the reputation, b_P in this case, when normalized to the domain $[0, 1]$ yields a utility $k \cdot b_P^2$, where k is the maximum benefit. A sublinear relationship between reputation and utility means that agents strongly favor those with high reputations. We use the relationship of $k \cdot \sqrt{b_P}$ for a superlinear relationship, which offers significant utility to all agents but those with the lowest reputations.

In our simplified interaction model, the agents alternate in granting favors. One can think of this as alternating delivery of an item by one agent, followed by payment by the other agent. With the linear

relationship between reputation and utility, a target with $b_P = 0.25$ would receive half the price for a good than would a target with a $b_P = 0.5$. The utility, u , of a target of type θ for a favor at time t , can be written simply as

$$u(p_B, t, \theta) = \gamma_\theta^t \cdot b_P \cdot \text{FAVORBENEFIT}. \quad (8)$$

3. Integrate Update and Payoff Functions. The update function and payoff function can now be integrated into Algorithms 1 and 2, where $n(r, \sigma_t)$ and $u(p_B(r), t, \theta)$ are used for `raterModel.GetNextReputation` and `raterModel.GetExpectedActionPayoff` respectively.

4. Run Algorithm 1 Over Domain of Reputations. In the basic Beta model, subsequent ratings affect an agent’s overall rating less than the previous. We examined a few different numbers of previous observations, but for the results reported in this paper, we used 10 previous observations. This means that we ran Algorithm 1 on each possible reputation with 10 observations, from 10 positive and 0 negative observations, through 0 positive reputations and 10 negative reputations (for other models we divided the reputation space into 10–100 points). Using other total numbers of observations to cover the full two dimensions of possible data is a valid approach, but we held the magnitude constant simply to rule out the Beta model’s nonstationarity.

Algorithm 1 also needs to be run with various discount factors for the strategic target agent. Except when otherwise noted, we ran discount factors from 0.0 to 0.8 in 0.1 increments.

Finally, the entire set of tests needs to be run with various values of FAVORBENEFIT and FAVORCOST to determine how consistently the model behaves across the range of favor sizes. For these values, we chose several combinations across the domains of c and b as outlined in Section 4.1.

5. Evaluate MONOTONICITY.

As our interaction model is focused on moral hazard, an IPS agent would prefer to interact with an agent with a higher discount factor than with an agent with a lower one. With a more patient agent, the IPS agent could enter into Nash equilibria in the repeated game that have higher payoffs for both agents. The IPS could do so using a trigger strategy, not unlike the related repeated prisoner’s dilemma model.

Given that the IPS agent prefers higher discount factors, we can examine whether more preferable strategic target agents have reputations that yield higher utility to raters in the interaction model. The payoff function maps the target’s reputation to its utility. The payoff functions for the Beta models are strictly monotonic (linear, square root, and quadratic). We evaluate these results with respect to the ranges of FAVORBENEFIT and FAVORCOST.

If the rater’s expected utilities are nondecreasing with respect to discount factor, then the reputation system is monotonic, as is the case with the Beta model with superlinear pricing. If the utilities are constant, as is the case with the Beta model with linear and sublinear pricing, then the reputation system is NONDISCRIMINATORY. If the rater utilities ever decrease with respect to increasing discount factor, then the system is nonmonotonic. Alternatively, if no meaningful asymptotic reputation exists, then the reputation system cannot be evaluated with respect to monotonicity.

6. Evaluate UNAMBIGUITY. We find UNAMBIGUITY by first examining each pair of successive inputs to Algorithm 1 for a given agent type (discount factor) and environment (FAVORBENEFIT and FAVORCOST). If the line defined by $r' = r$ is crossed between those two reputation values inclusively, then the point of intersection is a fixed point. If zero or multiple fixed points exist, then the system fails UNAMBIGUITY. Otherwise, we use this unique fixed point value of r when computing the other measures. We note that if insufficient resolution is used in evaluating possible input reputations with Algorithm 1 then

additional fixed points may be lost. For our results, we examined higher resolution outputs for subsets of our experimental results to make sure we were not likely missing any, though it is difficult to guarantee this numerically for reputation systems that exhibit noisy results.

7. Evaluate ACCURACY. After computing the fixed point to determine UNAMBIGUITY, it is relatively straightforward to calculate ACCURACY by computing the normalized mean absolute distance from each output of Algorithm 1 to the fixed point reputation for each agent type and environment.

8. Evaluate CONVERGENCE. Computing CONVERGENCE is also straightforward once the fixed point has been found. The slope may be closely approximated by computing the slope of the line segment between the points immediately surrounding the fixed point (or averaging the two nearby slopes if the fixed point lies on the boundary between two line segments).

4.4 Results

Here we discuss the results for each of the models we evaluate. We use our desiderata to compare reputation systems and find out how well they perform when faced with a strategic target agent. In doing so, we also validate that our desiderata are granular enough to distinguish differences between reputation systems, and that our results are intuitive.

Table 1 shows a summary of the reputation systems used and how they map into the model-specific functions for finding the next reputation and computing the expected action payoff. Table 2 shows a summary of our results discussed in the remainder of this section.

4.4.1 Beta Models Results

The Beta model, as described in Section 4.3, is the foundation for many approaches to reputation systems, including that proposed by Jøsang’s [1998; 2009] *Subjective* model, Teacy et al.’s [2006] *Travos* system, and Wang and Singh’s [2006; 2007] *Certainty* model. Most of the differentiation between these models is how they measure and aggregate uncertainty of reputation, but the underlying measurements are the same. We refer to this class of reputation systems as the *Beta* model.

Whereas the Beta models deal with the expected value of the probability that a target is trustworthy, many Beta models also focus on the uncertainty of this rating. This uncertainty is useful for determining whether to interact with a particular agent. Uncertainty can be an important element of decision-making for a risk-averse agent, that is, one who would prefer to avoid transactions that might have a negative outcome, even if the expected value is positive. To evaluate the effect of uncertainty as measured by the Travos and Certainty models, we reduce the utility expected from agents of uncertain trustworthiness. In the case of Travos, we multiply the expected utility by both the probability of a positive transaction and the certainty. For the Certainty model, we simply multiply the expected utility by the agent’s belief value, as this accounts for the both probability of a positive transaction and the uncertainty. In both models, certainty is in the range of $[0, 1]$.

The Beta model and the Subjective model exhibit nearly identical results, and so we examine them together. This is to be expected as the Subjective model’s belief is $\frac{\alpha}{\alpha+\beta+1}$. We did not examine the Subjective

Table 1: Summary of reputation systems evaluated.

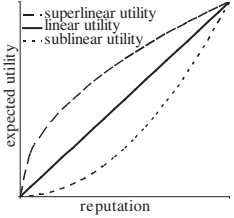
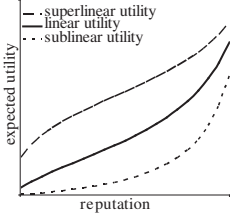
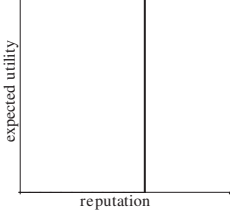
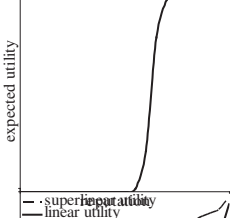
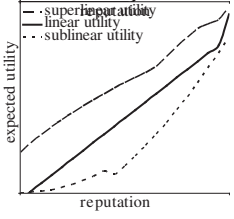
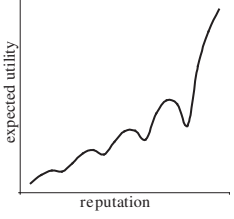
Reputation System	Next Reputation	Expected Action Payoff	Expected Action Payoff Graph
Beta	Increment the positive or negative experience count.	Exponentiate probability of positive outcome to 0.5, 1, or 2, for super-linear, linear, and sublinear, respectively, then multiply by favor value b . The graph shape is due to the straightforward expected value calculations.	
Certainty	Increment the positive or negative experience counts, compute the certainty of information.	Use Certainty model's belief in place of the Beta model's probability of a positive outcome. The graph curves are due to the additional factor of uncertainty.	
Discount Factor	Measure discount factor. Update probability distribution using Bayesian inference.	If the discount factor is sufficient to sustain full reciprocity then offer full favor. The graph is a step function produced by the cutoff value of the target's expected discount factor.	
Probabilistic Reciprocity	Add all accumulated favors to compute total balance.	Multiply the favor value by the probability of offering a favor. The graph shape is due to the model's sigmoid function.	
Sporas	Exponentially dampen old rating and combine with new rating.	Use the rating normalized to $[0, 1]$ in place of the Beta model's probability of a positive outcome. The discontinuities in the function's derivative arise due to points when the optimal strategy changes.	
Travos	Increment the positive or negative experience counts. Compute most probable bin in the Beta distribution.	Multiply the probability of a positive outcome by Travos's probability of being in the corresponding bin in the Beta distribution. Each non-monotonicity occurs when the reputation value is near the edge of a bin.	

Table 2: Summary of reputation system performances; the values listed are approximate averages across our experiments.

Reputation System	Unambiguity	Monotonicity	Convergence (slope, lower is better)	Accuracy (error, lower is better)
Beta (superlinear)	yes	monotonic	0 and 0.9	0.4
Beta (linear, sublinear)	yes	nondiscriminatory	0.9	0.45
Certainty	no	—	1	—
Discount Factor	yes	monotonic	< 0.1	0.02
Prob. Reciprocity	no	monotonic	no	0.2
Sporas (superlinear, linear)	yes	monotonic	≈ 0	0.3
Sporas (sublinear)	yes	nonmonotonic	no	0.4
Travos	yes	monotonic	0.8	0.2

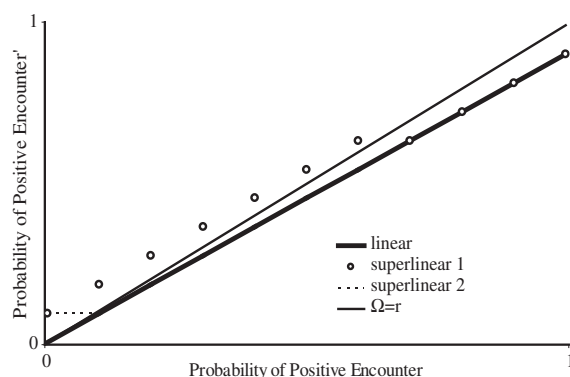


Figure 8: Beta and Subjective models.

model’s trust propagation, as it requires significant additional assumptions about beliefs of other agents’ digital signatures, which is not within our present scope.

The quality of the Beta model results varied by the interpretation of the probability of an agent performing positively. Using a linear interpolation of the probability, which is the natural risk-neutral way of modeling utility, led to results where no agents offered any favors and simply spent their reputations. The thick line in Figure 8 indicate typical results of such a linear probability-utility relationship, where all targets’ reputations converged toward the minimal reputation. The sublinear results were the same as the linear. The Beta model did not fare well on this case; the model fails MONOTONICITY, as all targets’ reputations end up the same. In the superlinear case, that is, where a target is either risk-seeking or is not harmed as much by negative interactions, the Beta model fares quite well. The superlinear Beta model meets CONVERGENCE with positive slopes, either slowly with slopes of 0.9 or at the ideal of 0, and also meets MONOTONICITY by distinguishing higher values of discount factors. The Beta model’s error in ACCURACY was mostly independent of the probability-utility relationship and ranged from 0.40 to 0.45.

The characteristics of the Certainty model became more pessimistic when evaluating against a group of three raters as opposed to an individual. The line labeled *network, probability* in Figure 9 shows the typical shape when a target is faced with a network of three raters. As shown by the lines labeled *individual, belief* and *individual, probability*, the targets were not incentivized to change their reputation until it crossed a critical threshold, at which point they would always perform positively. The Certainty model met neither

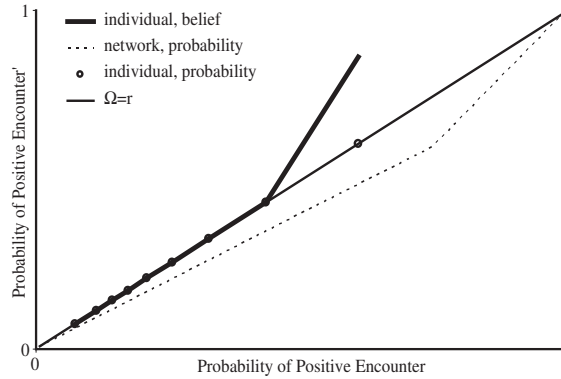


Figure 9: Certainty model.

UNAMBIGUITY nor MONOTONICITY, which made it difficult to assess CONVERGENCE and ACCURACY.

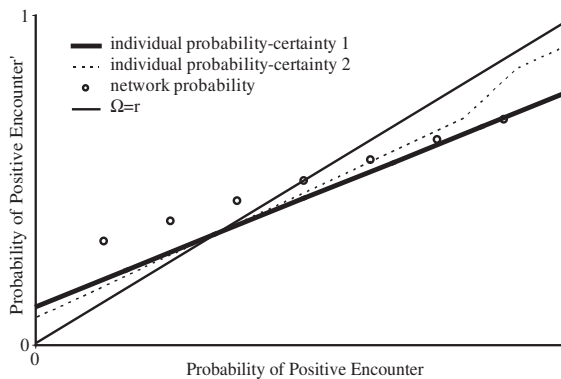


Figure 10: Travos model.

Travos computes uncertainty by subdividing the reputation space into five equal regions, finding the region containing expected probability of trustworthiness, and measuring certainty as the probability that the reputation is within the region. Travos normalizes the magnitude of all reputation information communicated to a rater to prevent one rater's recommendation from strongly dominating another rater's recommendation. However, this mechanism also amplifies small numbers of observations, as the aggregation mechanism implicitly assumes a relatively large number of observations. Travos did not meet MONOTONICITY, as all of the parameterizations yielded the same fixed point, which may be due to the normalization, requiring a significant volume of transactions to change the fixed point. Given that all of the reputations converged to the same point, the fact that Travos generally met the other desiderata does not carry significant weight in evaluating the model.

4.4.2 Probabilistic Reciprocity Results

Sen [2002] proposed the *Probabilistic Reciprocity* model as a way for an agent to experiment with trusting another agent to see if the first agent reciprocates favors back. Each agent keeps track of the total amount of utility spent and gained throughout the history of games between itself and other agents, summing the

utilities of the gains and losses as the balance, B . Agents use this balance to adjust their probability of performing a favor to another agent. The probability function is written in terms of the cost of the current favor, c , the expected cost to offer a favor, $E(C)$, as

$$P(\text{offer favor}|B) = 1 / \left(1 + \exp\left(\frac{c - \beta E(C) - B}{\tau}\right) \right). \quad (9)$$

The parameters β and τ are tunable cooperation constants. We use balance, B , as an agent's reputation, as this is the only parameter that encodes reputation information.

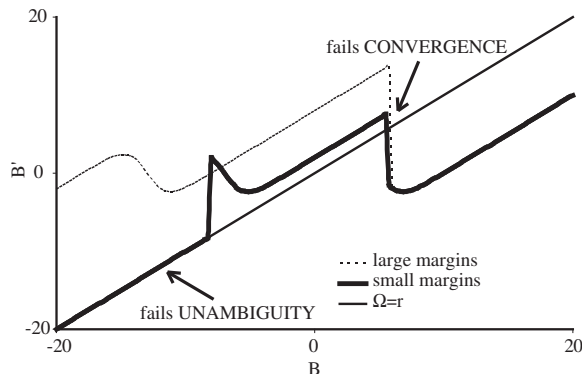


Figure 11: Probabilistic Reciprocity model.

The Probabilistic Reciprocity model, depicted in Figure 11, meets some of the desiderata most of the time, but does not converge. This figure shows two examples, one where the benefit of the favors is significantly larger than the costs (large margins: $c = 10$, $b = 18$), and one where the benefit is only slightly larger than the cost (small margins: $c = 10$, $b = 12$). The model generally met MONOTONICITY in every occurrence we examined, excluding ranges of fixed points where an agent's initial reputation is too low, such as the left portion of the line labeled *small margins*. Figure 11 shows such a range in the lower portion of the thick line. In these cases, the model fails UNAMBIGUITY because agents would refuse to consider dealing with an agent with a reputation that is too low, leaving its reputation unchanged. The weakest part of the model was CONVERGENCE, as the magnitude of the slope at the fixed point, $|\frac{d\Omega}{dr}|$, was far greater than 1 in all cases, and always negative. This means that an agent's reputation often changes significantly after every successive interaction and never converges. Finally, across our various parameterizations, the ACCURACY of the model was usually around 0.2, but was as low as 0.11 and as high as 0.22. The model's error was lowest when parameterized at moderate to large margins, such as $c = 10$ and $b = 18$, as opposed to those with highest or lowest margins (such as either $c = 10$ and $b = 12$, or $c = 10$ and $b = 30$).

4.4.3 Discount Factor Results

Hazard [2008] and Smith and desJardins [2009] both proposed variations of the *Discount Factor* model, in which agents strategically maximize utility while attempting to discover each others' discount factors. An agent's discount factor is a measure of the agent's patience, weighting how the agent accounts for future utility by an exponentially decreasing function of time. In this model, the expected value of an agent's discount factor is its reputation. An agent with a discount factor close to 0 would be myopic and greedy, whereas an agent with a discount factor close to 1 would offer favors if it expects the relationship or global

reputation from offering a favor to be beneficial to itself in the long run. Like the Probabilistic Reciprocity model, the reputation of the Discount Factor model is explicitly connected with agents' utilities.

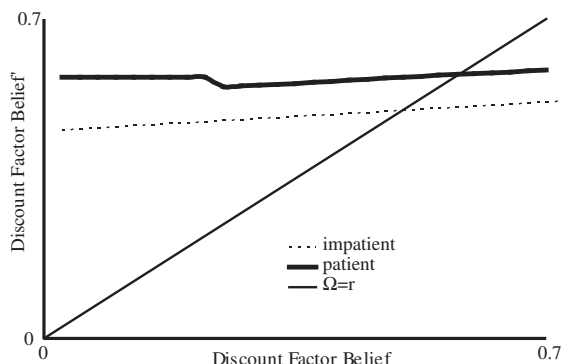


Figure 12: Discount Factor model.

Figure 12 shows the results of the Discount Factor model. Across all the parameterizations we examined, the results were similar to this graph with all lines of the same shape, the only major variation being the vertical location of the line on the graph. Because the agents in the model are strategic, they choose the optimal strategy that corresponds to their discount factors. Targets cannot credibly maintain an incorrect reputation, and their reputations converge quickly. We found that agents with a higher discount factor always offer better utility to a patient agent, so MONOTONICITY is met. Each agent type also had exactly one fixed point, so UNAMBIGUITY is also met. The model fared well with the CONVERGENCE desideratum, with $\frac{d\Omega}{dr}$ being small and positive, usually less than 0.1. The error was small, and so this model performed well with regard to ACCURACY. Across all our parameterizations, the error was between 0.014 and 0.028.

4.4.4 Sporas Results

Zacharia and Maes [2000] propose the *Sporas* reputation model which measures targets' reputations according to a specified range, with the rater's reputation influencing the magnitude of the reputation change. This model employs a dampening function that slows the maximum rate at which a target's reputation may change for a given observation as the target's reputation increases. Zacharia and Maes motivate Sporas based on online marketplaces and use continuous reputation values.

The Sporas model behavior, depicted in Figure 13, was remarkably similar to that of the Beta model, even though the Sporas model permitted continuous interactions. Although we were initially surprised at the similarity, both models' reputation computations have a linear term of quality. Sporas model did not meet CONVERGENCE in the sublinear case or when the difference between c and b was large. The error in ACCURACY for Sporas was slightly better than the Beta model, ranging from 0.20 to 0.45.

5 Discussion

The primary purpose of a reputation system is to provide information to agents about other agents with the goal of improving social welfare. This goal assumes that if agents know which other agents are trustworthy and which agents are likely to defraud, then the agents' utilities would be improved as agents would self-select transactions with preference toward trustworthy agents. Whereas a possible design goal is to increase

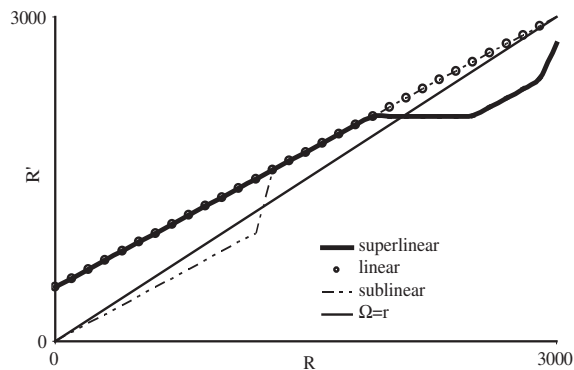


Figure 13: Sporas model.

trustworthy agents' utility the most, the primary goal is to inform agents and reduce uncertainty with regard to some interaction mechanism.

If an interaction mechanism has the property of incentive compatibility, then all strategic targets would always play honestly according to their valuations. Although incentive compatible systems may have use for multiagent learning, for example to determine which agents might receive the most benefit from which products, such systems have little need for an additional system to measure the reputations of targets. If a reputation measurement system were added in order to measure targets' reputations for use in an additional context or situation, incentive compatibility may no longer hold. Whereas our desiderata would work in measuring reputation in an incentive compatible mechanism, the measurements may have reduced relevance. Alternatively, if the agents in a system exhibit highly specific behavior and are not strategic, then our desiderata would need to be modified to use the specific behavior in place of the strategic behavior. Such a system can arise when interactions are only permitted via proxy agents, that is, targets that have a predefined behavior that act based on a specific set of parameters.

Our desiderata are useful measures for how well a reputation system will hold up against strategic attacks. For example, Kerr and Cohen [2009] outline a number of possible ways that an agent could strategically improve its utility by being dishonest in a reputation system. Their "reputation lag attack," achieved by a target alternating between honest and cheating periods, is applicable when a reputation system that fails to meet CONVERGENCE because a target can exploit oscillations of its reputation. Similarly, their "value imbalance attack," achieved by a target being honest with low-cost goods and dishonest with high-cost goods, and "reentry attack," where an agent continually opens new accounts to dishonestly use a new untainted reputation, both indicate that a reputation system has poor ACCURACY. A reputation system designed for high ACCURACY would recognize dishonest targets more quickly.

When analyzing pure moral hazard situations, the resulting Nash equilibria are often mixed strategies, where a target chooses its actions stochastically based on some distribution. If mixed strategies are necessary or desirable for a particular reputation mechanism, then the reputation system should somehow recognize when a target is employing a mixed strategy. Detecting whether a mixed strategy is being employed has some uncertainty to it, as it must be done statistically within some bounds of confidence. If the reputation system does not collapse a mixed strategy into a single reputation value, a system may fail CONVERGENCE over some range of behavior even if it is an effective system.

One difficulty in evaluating a reputation measure is if the measurements are nonstationary, meaning that the reputation measures themselves somehow change over time. Nonstationarity can arise if it becomes

increasingly more difficult or easy to change a reputation when further measurements are made, as is the case when interactions are aggregated over the entire lifetime of a target without any sort of dampening—i.e., without weighting old interactions less than recent ones. Whereas some reputation systems, such as Amazon Marketplace, Travos, and Certainty, employ nonstationary measures, such reputation systems must be used with caution because the difficulty of a target changing its reputation becomes increasingly difficult as a function of the target’s age, as even the oldest interactions count as much as recent ones.

Our desiderata do not always indicate that one reputation system is the best one for a particular situation. The choice of which reputation system to employ comes down to trade-offs. For example, one system may offer better CONVERGENCE whereas another may offer better ACCURACY. Having good CONVERGENCE means that the given system quickly reaches an equilibrium where the reputation is close to the actual value, but this can be misleading in cases when the reputation dynamics change rapidly close to the fixed point. A system’s having good ACCURACY means that it corrects a target’s reputation to achieve a reasonably accurate value quickly, even if the initial reputation is far off. However, raters may be able to only discern a small amount of information from each transaction in some interaction mechanisms, and so the interaction model may be detrimental to ACCURACY. If a system does not exhibit UNAMBIGUITY, but the unreasonable fixed points are impossible to reach by the path a target’s reputation takes, then those unreasonable fixed points may be ignored. However, if unforeseen events, such as a software glitch, incorrectly push a target’s reputation into these regions, then the ignored fixed points become extremely important and can possibly have major negative impacts to the reputation system as a whole.

Reputation systems may work better in one domain than another. A reputation system may work well in the case of adverse selection, but perform poorly in the case of moral hazard. The effectiveness of a reputation system may change drastically even with different parameters in the same environment, even if the only different parameter is the topology of agent relationships. Therefore, when applying our desiderata to a reputation system, they should be applied to a setting as close to the actual environment as possible. If parameters of the environment or interactions are known to change quickly or drastically, then the desiderata should be employed across the range of environments and interactions. One reputation system may perform well in a certain niche case, but may perform poorly across the full range of interactions.

6 Conclusions and Future Work

The four desiderata we present measure how well a mechanism fares at measuring the reputation of a strategic agent. Of the systems we examined, the one that takes moral hazard as the primary environment in designing the system, the Discount Factor model, fared the best when evaluated in a moral hazard situation. The results from applying our desiderata were granular enough to differentiate reputation systems.

We focus on strategic agents because they maximize their own utility and are thus generally more attractive to users. For example, a business would tune the decision models in one of its webservice products to maximize profit, or a user of a peer-to-peer file sharing service might attempt to change the peer-to-peer client software to achieve faster downloads. An agent with a high reputation may have considerably greater ability to cause harm to other agents than an agent with a low reputation, enabling a malicious agent to strategically build up reputation to maximize the harm it causes. A variation on our measures would be to use a strategically malicious agent, whose utility is a function of the loss of other agents. Many applications of reputation systems involving businesses and consumers, particularly those where an autonomous agent is acting on behalf of the firm or individual, will be faced against rational agents. However, a strong case may be made for modeling with strategically malicious agents for use in social networks, in businesses that

might expect malice from extortionists or angry customers or competitors, or in using reputation as a basis for finding and tracking terrorists.

Collusion, side-payments, and Sybil attacks (using many pseudonyms to boost or reset reputation) are other exceptions when agents may appear to not act individually rational. However, our desiderata can be adapted to measure the reputation dynamics given a certain number of colluding raters attempting to boost the reputation of one scamming agent. To extend these desiderata, the colluding raters should be treated as one agent in terms of utility. The reputation of the scamming rater, that is, the colluding agent with the reputation inflated by the other colluding raters, can then be used directly in the desiderata.

The biggest weakness of our desiderata list is the computational complexity required to model reputation aggregation across a large number of agents and against strategic agents with high discount factors. Because we simply exhaust all possible actions, a number of states exponential in the number of actions must be computed. Solving a specific reputation system behavior against a strategic agent may be feasible with a simple reputation system and lead to an efficient solution, but large and complex reputation systems, particularly those without closed form solutions and highly domain-specific features (i.e., those having complex relationships between the reputation system and the interaction model), exacerbate the matter. Graphing Ω can offer insight into the dynamics of a reputation system, but visualizing Ω may be nontrivial for systems that employ reputations of high dimensions that do not collapse easily to a scalar value. Determining measurable desiderata that work well in complex scenarios is an interesting and useful avenue for future work.

Because our desiderata are measured against a rational agent that would take advantage of any weaknesses of the reputation system, obtaining conclusive results for a reputation system intended for human involvement requires a sizeable controlled experiment. Such empirical results would need to deal with the significant noise in the system and would require sufficient data to conclude that a fixed point of a reputation is a stationary process. Evaluating currently deployed systems with respect to our desiderata, although a significant undertaking, would contribute both to the understanding of our desiderata and the reputation systems.

Whereas our desiderata are useful for ensuring that reputation systems are useful to the agents, strictly following the desiderata is not always in the best interest of the party that implements the marketplace or mechanism—as opposed to the agents who interact with each other in context of the marketplace. A firm setting up an online auction that profits from each transaction has an incentive to maximize targets' reputations, that is, maximize Ω , such that agents perform transactions before the agents realize that not all are trustworthy. However, such a practice is not sustainable, and so a firm looking for long-term profits would need to ensure that the reputation system is useful. Such a firm would need to make trade-offs among short-term profit, long-term profit, and the various desiderata.

Our desiderata are by no means exhaustive and may be modified or extended for domain-specific purposes. They make a good start toward a general framework for directly comparing the effectiveness of different reputation systems in a specified situation.

Acknowledgments

We are indebted to the anonymous reviewers whose comments helped greatly improve our results and presentation.

References

- Altman, A., Tennenholtz, M., 2008. Axiomatic foundations for ranking systems. *Journal of Artificial Intelligence Research* 31, 473–495.
- Axelrod, R., July 2000. On six advances in cooperation theory. *Analyse & Kritik* 22, 130–151.
- Barber, K. S., Fullam, K., Kim, J., 2003. Challenges for Trust, Fraud and Deception Research in Multi-agent Systems. Vol. 2631. Springer Berlin / Heidelberg, Ch. 2, pp. 167–174.
- Dellarocas, C., June 2005. Reputation mechanism design in online trading environments with pure moral hazard. *Information Systems Research* 16 (2), 209–230.
- Devaney, R. L., October 1992. *A First Course in Chaotic Dynamical Systems: Theory and Experiment*. Westview Press, Boulder, Colorado.
- Dingledine, R., Freedman, M. J., Molnar, D., February 2000. Accountability measures for peer-to-peer systems. In: Oram, A. (Ed.), *Peer-to-Peer: Harnessing the Power of Disruptive Technologies*. O'Reilly, Ch. 16, pp. 271–340.
- Ely, J. C., Välimäkiz, J., March 2003. Bad reputation. *The Quarterly Journal of Economics* 118 (3), 785–814.
- Fudenberg, D., Levine, D. K., 1992. Maintaining a reputation when strategies are imperfectly observed. *The Review of Economic Studies* 59 (3), 561–579.
- Fullam, K. K., Barber, K. S., May 2006. Learning trust strategies in reputation exchange networks. In: *Proceedings of the 5th International Conference on Autonomous Agents and Multiagent Systems*. Hakodate, Japan, pp. 1241–1248.
- Fullam, K. K., Klos, T. B., Muller, G., Sabater, J., Schlosser, A., Topol, Z., Barber, K. S., Rosenschein, J. S., Vercouter, L., Voss, M., 2005. A specification of the agent reputation and trust (ART) testbed: experimentation and competition for trust in agent societies. In: *Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems*. ACM, New York, NY, USA, pp. 512–518.
- Goffman, C., Pedrick, G., 1983. *First Course in Functional Analysis*. AMS Bookstore, Ch. 2, p. 71.
- Hazard, C. J., July 2008. ¿Por favor? Favor reciprocation when agents have private discounting. In: *AAAI Workshop on Coordination, Organizations, Institutions and Norms (COIN)*. Chicago, Illinois, pp. 9–16.
- Hoffman, K., Zage, D., Nita-Rotaru, C., December 2009. A survey of attack and defense techniques for reputation systems. *ACM Computing Surveys* 41 (4), to appear.
- Houser, D., Wooders, J., 2006. Reputation in auctions: Theory and evidence from eBay. *Journal of Economics & Management Strategy* 15 (2), 353–369.
- Huynh, T. D., Jennings, N. R., Shadbolt, N. R., 2006. An integrated trust and reputation model for open multi-agent systems. *Journal of Autonomous Agents and MultiAgent Systems* 13 (2), 119–154.

- Jøsang, A., 1998. A subjective metric of authentication. In: Proceedings of the 5th European Symposium on Research in Computer Security. Springer-Verlag, London, UK, pp. 329–344.
- Jøsang, A., Ismail, R., Boyd, C., March 2007. A survey of trust and reputation systems for online service provision. *Decision Support Systems* 43 (2), 618–644.
- Jøsang, A., Quattrociocchi, W., 2009. Advanced features in Bayesian reputation systems. In: *Trust, Privacy and Security in Digital Business*. Vol. 5695 of Lecture Notes in Computer Science. Springer Berlin / Heidelberg, pp. 105–114.
- Jurca, R., Faltings, B., August 2007. Obtaining reliable feedback for sanctioning reputation mechanisms. *Journal of Artificial Intelligence Research* 29, 391–419.
- Kamvar, S. D., Schlosser, M. T., Garcia-Molina, H., 2003. The eigentrust algorithm for reputation management in P2P networks. In: Proceedings of the 12th International Conference on World Wide Web. Budapest, Hungary, pp. 640–651.
- Kerr, R., Cohen, R., 2006. Modeling trust using transactional, numerical units. In: Proceedings of the 2006 International Conference on Privacy, Security and Trust. ACM, New York, NY, USA, pp. 1–11.
- Kerr, R., Cohen, R., 2009. Smart cheaters do prosper: Defeating trust and reputation systems. In: Proceedings of the eighth international conference on autonomous agents and multiagent systems. pp. 993–1000.
- Krupa, Y., Hubner, J. F., Vercouter, L., 2009. Extending the comparison efficiency of the ART testbed. In: Proceedings of the First International Conference on Reputation - Theory and Technology. Gargonza, Italy, pp. 186–199.
- Melnik, M. I., Alm, J., 2003. Does a seller's eCommerce reputation matter? Evidence from eBay auctions. *The Journal of Industrial Economics* 50 (3), 337–349.
- Ramchurn, S. D., Huynh, D., Jennings, N. R., 2004. Trust in multi-agent systems. *The Knowledge Engineering Review* 19, 1–25.
- Resnick, P., Kuwabara, K., Zeckhauser, R., Friedman, E., 2000. Reputation systems. *Communications of the ACM* 43 (12), 45–48.
- Resnick, P., Sami, R., 2007. The influence limiter: Provably manipulation resistant recommender systems. In: Proceedings of the ACM Conference on Recommender Systems. Minneapolis, MN, pp. 25–32.
- Resnick, P., Sami, R., 2008. The information cost of manipulationresistance in recommender systems. In: Proceedings of the ACM Conference on Recommender Systems. Lausanne, Switzerland, pp. 147–154.
- Resnick, P., Zeckhauser, R., Swanson, J., Lockwood, K., 2006. The value of reputation on eBay: A controlled experiment. *Experimental Economics* 9 (2), 79–101.
- Saha, S., Sen, S., Dutta, P. S., July 2003. Helping based on future expectations. In: Proceedings of the Second International Joint Conference on Autonomous Agents and Multiagent Systems. Melbourne, Australia, pp. 289–296.
- Salehi-Abari, A., White, T., 2009. Towards con-resistant trust models for distributed agent systems. In: International Joint Conference on Artificial Intelligence. pp. 272–277.

- Sen, S., December 2002. Believing others: Pros and cons. *Artificial Intelligence* 142 (2), 179–203.
- Sen, S., Banerjee, D., 2006. Monopolizing markets by exploiting trust. In: *Proceedings of the Fifth International Joint Conference on Autonomous Agents and Multiagent Systems*. Hakodate, Hokkaido, Japan, pp. 1249–1256.
- Sen, S., Goswami, I., Airiau, S., 2006. Expertise and trustbased formation of effective coalitions: an evaluation of the art testbed. In: *Workshop on Trust in Agent Societies at The Fifth International Joint Conference on Autonomous Agents and Multiagent Systems*. pp. 71–78.
- Sierra, C., Debenham, J., 2005. An information-based model for trust. In: *Proceedings of the 4th International Joint Conference on Autonomous Agents and Multiagent Systems*. ACM, New York, NY, USA, pp. 497–504.
- Smith, M. J., desJardins, M., February 2009. Learning to trust in the competence and commitment of agents. *Autonomous Agents and Multi-Agent Systems* 18 (1), 36–82.
- Sonnek, J. D., Weissman, J. B., 2005. A quantitative comparison of reputation systems in the grid. In: *Proceedings of the 6th IEEE/ACM International Workshop on Grid Computing*. pp. 242–249.
- Srivatsa, M., Xiong, L., Liu, L., 2005. Trustguard: Countering vulnerabilities in reputation management for decentralized overlay networks. In: *Proceedings of the 14th International Conference on World Wide Web*. pp. 422–431.
- Teacy, W. T., Patel, J., Jennings, N. R., Luck, M., 2006. Travos: Trust and reputation in the context of inaccurate information sources. *Autonomous Agents and Multi-Agent Systems* 12 (2), 183–198.
- Wang, Y., Singh, M. P., July 2006. Trust representation and aggregation in a distributed agent system. In: *Proceedings of the 21st National Conference on Artificial Intelligence*. pp. 1425–1430.
- Wang, Y., Singh, M. P., 2007. Formal trust model for multiagent systems. In: *Proceedings of the 20th International Joint Conference on Artificial Intelligence*. Hyderabad, India, pp. 1551–1556.
- Zacharia, G., Maes, P., October 2000. Trust management through reputation mechanisms. *Applied Artificial Intelligence* 14 (9), 881–907.
- Zhang, J., Cohen, R., 2007. A comprehensive approach for sharing semantic web trust ratings. *Computational Intelligence* 23 (3), 302–319.