# From subjective reputation to verifiable experiences — augmenting peer-control mechanisms for open service ecosystems

Sini Ruohomaa, Puneet Kaur, and Lea Kutvonen

University of Helsinki, Finland
sini.ruohomaa@cs.helsinki.fi,
WWW home page: http://cinco.cs.helsinki.fi/

**Abstract.** In inter-enterprise collaborations, autonomous services from different organizations must independently determine which other services they can rely on. Reputation-based trust management in Pilarcos utilizes shared experience information on the actors' past behaviour in estimating the risks of a collaboration; these experiences are shared between members of the service ecosystem through a reputation system. As the reputation system becomes an essential peer-control mechanism for the open service ecosystem, it must be augmented with sanctions for misbehaviour and appropriate incentives for correct behaviour. A fair sanctioning system cannot be built on traditional subjective reports, as rebuttal of undeserved reports requires shared, objective measures. To make the shared experience information objective and verifiable, we associate it with whether the relevant collaboration contract was followed, backed up with evidence in the form of nonrepudiable receipts. In this way, we are able to protect automated reputation-based trust decisions from being skewed by misinformation.

## 1 Introduction

In inter-enterprise collaborations, services from different organizations and domains join together to fulfil a mutual goal. In the open service ecosystem, the services are autonomous, and there is no centralized control of the collaboration process. Each service must independently determine which other services it should collaborate with; a trust decision is made to determine this willingness to rely on another service. The Pilarcos inter-enterprise collaboration management infrastructure [9, 8] contains a trust management system to automate these decisions in routine cases [17]; selected difficult or high-stake decisions are forwarded to a human user based on policy-defined rules [6].

A central element of the trust decisions is an estimation of the risk of collaborating with the given actor. The estimate is based on gathered experiences on the actor's past behaviour, which consists of both first-hand experiences from earlier collaborations with it, and shared experiences from third parties received through a reputation system.

A reputation system has two tasks. First, from the enterprise perspective, it provides information to support trust decisions for each individual service; this helps the participants in the reputation system to find new partners and to steer clear of misbehaving services in order to limit their risk.

Second, from the ecosystem perspective, it introduces a form of peer control where misbehaviour towards other actors or the ecosystem infrastructure is punished through reputation loss. Sociological research indicates that direct first-order punishment for misbehaviour is not enough by itself to ensure that the ecosystem can scale up in size: it must be complemented with a second-order punishment to discourage unfair first-order punishments [3]. In other words, spreading false experiences must have a negative reputation impact on the source.

To contrast these requirements to current solutions, systems sharing reputation information are occasionally also considered as subjective recommender systems on other users; with such an approach, the aim is to promote commonly liked services rather than implement robust peer-based control. For a recommendation system, experiences are accepted as subjective reports on the fulfilment of expectations. As different expectations can lead to different reports even on identical behaviour, these reports do not objectively describe actual outcomes in the sense that reports on breaches of contract do. While they can still support individual decision-making as indicators of popularity, the subjectivity of the criteria in use makes recommender systems unsuitable for social control. Solutions have been proposed to promote similar understanding of the recommendation values [5, 2]; however, shared semantics do not yet change the fact that the actors involved are stating their opinions, which cannot be used as a basis of judging whether a statement is unfair. As an extreme example, two honest actors may judge a musician's performance completely differently based on their tastes, due to a lack of objective measurement scale for what is "good music". To get around this ambiguity, we propose to monitor for and report objectively defined events: explicit breaches of collaboration contracts.

The core problems of unverifiability and subjectivity of experiences in current approaches hinder the use of reputation information in inter-enterprise collaborations, particularly for automated decision-making. Falsely accused actors must be able to rebut reports to clear their name, while honest reporters should be protected from retaliatory action, and the reciprocity of feedback [16] limited.

We advance a reputation system based on objective, verifiable experiences. It is designed to support automated trust decisions on inter-enterprise collaborations, and implements peer control in the service ecosystem by supporting second-order punishment: a successful rebuttal of a false experience causes reputation loss for the dishonest information source. This is achieved by introducing a nonrepudiable audit trail to collaborations, and defining reputation impacts of misbehaviour in collaboration contracts. Similarly, negative reputation impacts for misreporting are defined in a reputation network contract. The solution extends the existing Pilarcos collaboration management infrastructure [9, 17].

The rest of the paper is structured as follows: Section 2 presents the Pilarcos ecosystem we build on, and related work. Section 3 maps evidence to reputation

impacts through collaboration contracts to achieve objectivity, and specifies the process of creating new reputation information in which unverifiable experiences can be rebutted and removed from the system. Section 4 discusses the impact of the solution and compares it to the state of the art.

## 2 Background and related work

In the first subsection, we summarize the existing Pilarcos open service ecosystem, in which we utilize reputation for trust management. The second subsection presents related work on the topic of objective and verifiable experiences.

### 2.1 Reputation-based trust management in Pilarcos

The Pilarcos collaboration management infrastructure provides support for partner discovery, interoperability management, contract negotiations, runtime monitoring, including contract breach detection and recovery, as well as local trust decisions evaluating the actors' willingness to collaborate with their potential partners [9, 17]. The Pilarcos service ecosystem is collaboratively governed, rather than centrally controlled, to ensure its long-term viability and scalability [19].

We propose to strengthen the implemented reputation-based trust management system in Pilarcos [17] by providing it with a flow of objective and verifiable reputation information. The trust management system is modular, and can take advantage of different kinds of reputation systems as its information sources. It splits experience information into four dimensions: monetary, reputation, control and satisfaction. This allows risk evaluations to differentiate between e.g. misbehaviour that directly causes monetary loss, deterioration of own reputation caused by a partner spreading fraudulent reports afterwards, weakening of peer control due to an actor's misbehaviour as a recommender, and failures to satisfy the demands set in contracts, which may or may not have direct monetary consequences, respectively [17].

Collaboration contracts, or eContracts, are based on business network models that specify the structure and business processes of the collaboration, and relevant trust decision points [8, 17]. These models are modular, reusable and public, and they are produced by domain experts in response to the needs of the ecosystem. During contract negotiations, open options in the models, such as particular quality of service requirements or the price of the service, can be further refined to form an agreement between the participants [7].

The trust management process can be divided into two parts: the trust decisions, and the evolution process of the reputation information. Both are governed by their own policies. A trust decision is triggered at specific points of the collaboration process, as further resources are committed and an up-to-date risk evaluation is needed. To evaluate the risk of proceeding with the collaboration, we predict the outcome it would have on different assets based on previous experiences, which are stored as reputation information; the details of the format are described in earlier work [17]. In this paper, we focus on the evolution of reputation information through sharing experiences in a reputation system.

### 2.2 Related work to support objective and verifiable experiences

We distinguish three approaches for collecting experiences in a way that members in the service ecosystem can agree on their content: centrally orchestrated, fully distributed and a protocol-based approach utilizing third-party witnesses.

TrustCoM [22] represents the centrally orchestrated approach. In TrustCoM, performance monitors both internal and external to the actors collect information pertinent to fulfilling the Service Level Agreement (SLA), such as response times. The monitors send these raw observations to an SLA Evaluator, which is a third party trusted to pass a neutral judgement on the transacting parties. This result is, in turn, reported to a trusted third party reputation system, or used as a basis of removing a partner from the collaboration. The approach follows a tradition set by centralized workflow execution, such as implemented by CrossWork [12]. It involves a trusted infrastructure service or a hub member of the ecosystem running a distributed business process by using the other participants as components. This central operator can judge the performance of the other actors, and decide on sanctions directly. The main difference in operational environments is that in Pilarcos, control and monitoring are distributed among the autonomous and not fully trusted participants. As a result, there are no actors that are able to observe all collaborations in the ecosystem.

In the fully distributed approach (e.g. TrustGuard [20]), the transacting parties exchange nonrepudiable, i.e. cryptographically signed, receipts that act as evidence of their actions later. The main challenge in this approach is in the asymmetry of receipts: the party responsible for signing the last receipt can refuse to finish the protocol, which leaves a hole in the audit trail of evidence [13]. In TrustGuard, receipts document the intention to collaborate, which means that the service is not provided before the protocol has been completed. This scheme protects against submitting experiences of transactions that were never really started, but cannot stop unfair reports from being made after the transaction.

In the protocol-based approach (e.g. Li, Martin and Zhang [11]), a third party witness is included in transactions, observing them and implementing a fair, "atomic" exchange [13]. The proposal of Li et al. aims to secure an electronic marketplace, which can host a simple form of inter-enterprise collaboration where goods are exchanged in brief, fixed transactions. It involves an arbitrator, who acts as a witness and judge of a collaboration, and punishes misbehaviour through withdrawals on a monetary deposit that all participants have made beforehand in a trusted bank. In addition, the arbitrator can report experiences to a reputation system. When the arbitrator is not needed to resolve a disagreement, a central broker service reports a default positive experience once a specified time has passed since it matched the two actors [11].

Li et al. implement a centralized punishment system on top of the fair receipt exchange protocol, which is where they differ from Pilarcos. Our aim in this paper is to distribute punishment as peer-based control, while taking advantage of the third party witness approach to provide a stream of verifiable evidence on the outcomes of transactions in the collaborations.

To distinguish between a passive witness and the active arbitrator passing judgement, we denote the former as *notaries*: a notary's task is to verify with its own signature whether a protocol was followed according to the specification it has been given by the transacting partners. The notary must be trusted by the given partners to remain impartial on that exchange to eliminate asymmetry, and their agreement on a given notary must be nonrepudiable once the transaction begins. On the other hand, a specific notary does not necessarily need to be trusted by anyone else in the collaboration or the entire ecosystem, as each independent exchange can be observed by a different notary. This limits the power that any specific notary service can gain over the actors in the marketplace.

As a related branch, certification-based trust also relies on cryptographical verifiability, but should not be confused with the proposed approach of signed receipts. Certification-based trust is used in e.g. NICE [10]. Instead of being experience-based, the system relies on signed expressions of "I, Alice, trust Bob", and trust decisions are based on policies on whether a provided set of trust declarations, or certificates, are sufficient to make the decision-maker also trust the considered target. Other examples of certification-based trust used for access control include WS-TRUST [14] and KeyNote [1]. Misinformation is a relatively minor issue for these systems, as the group of accepted information sources is small, closed and managed offline: certificates document networks of pre-formed trust relationships between the sources rather than guiding their evolution.

In summary, related work shows that signed, nonrepudiable receipts provide a promising basis for the verifiability of experiences in open service ecosystems. We have found that third party witnesses, notaries, are needed to guarantee the fair exchange of receipts on the transaction, and that the concept has already been applied within electronic markets.

In the following section, we apply the solution to provide a basis for objective and verifiable reputation information in the open service ecosystem by first providing a mapping between the receipt evidence and the corresponding experience stored in a reputation system through contracts, then specifying a process for creation of new reputation information that allows experiences not backed by appropriate evidence to be rebutted and removed from the reputation system.

## 3 Sharing and rebuttal of new experiences

In order to make reputation information objective, it must be defined so that it has a measurable truth value, rather than as a subjective opinion. In this section, we define the basis for objective experiences, and specify the process for submitting them into the reputation system as well as for rebuting false experiences.

For inter-enterprise collaborations, objective experience sharing is made feasible by electronic contracts governing the collaborations: if the contract was followed, experiences should be positive; if it was violated, they should be negative. From the perspective of a single service, the impact of a specific outcome may vary from minor to major gain or loss [17], but in order to make the shared

experience objective, the impact in the shared experience must be standardized through the contract as well. A natural source for this information is the business network model referenced by the contract. The business network model is a formal model for the collaboration and defines e.g. the communication protocols and compensation processes involved [9].

Determining the reputation effect of different outcomes becomes a part of the modelling process done by the domain expert, which makes the mapping reusable over multiple collaboration instances using the same model. The modelled values can be further fine-tuned in the contract negotiations, in case the same model can be used for collaborations dealing with very different stakes.

The collaboration contract, then, should specify a mapping of outcomes to experiences, for example that the event "goods received" should translate to an experience with major positive effect on the monetary and satisfaction assets, and no effect on the control and reputation assets [17]. In this example, the party receiving the goods (A) will submit this experience on the party who sent them (B). In order to support verifiability, it also produces a signed, non-repudiable receipt of the outcome of the step. The example communication protocol and the following experience submission process are depicted in Figure 1.

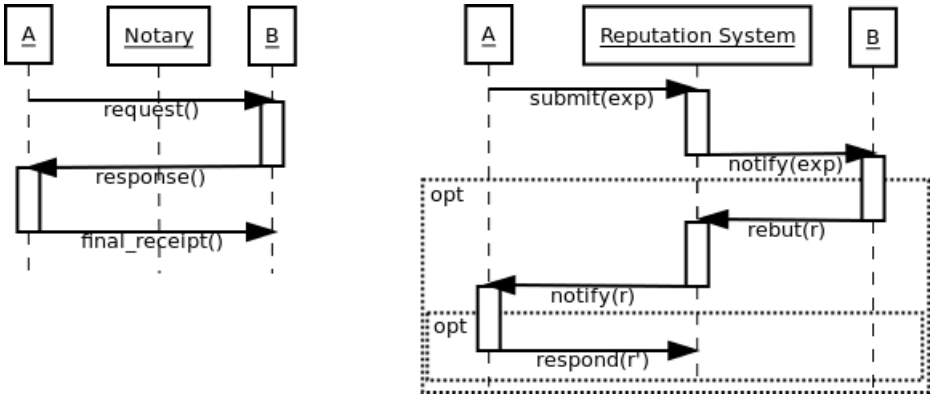


**Fig. 1.** An example collaboration protocol and submission of an experience report.

The receipts, depicted in Figure 2, remain private unless they are used to rebut an experience, and they are stored by both collaborators. The source actor (A) is the provider of the receipt. The target actor (B) is the actor for whom the receipt is generated. The notary witnesses the transaction step when the protocol in the collaboration model so defines. In the optimistic fair exchange protocol [13], collaborators can choose to only involve the notary if one party fails to respond on time. This requires that the relevant exchange can be repeated with the notary listening in as needed, but it should significantly reduce the communication overhead: the threat of notary involvement removes the attractiveness of receipt omissions almost as effectively as the observation itself [13].

- Source actor id, target actor id
- Notary id
- Business network model id, eContract id, task id
- Task counter, receipt protocol step
- Outcome of step (e.g. "goods received")
- Signature of witness (source or notary)

**Fig. 2.** The contents of a receipt.

The model and task information together provide a reference for determining the reputation impact represented by the receipt, and a unique identifier for the receipt (and corresponding experience) in case the same task is repeated multiple times. Each task can be expected to produce several receipts on relevant steps, most notably its start and end, and this is captured by the step identifier. The outcome of the step is the identifier of the given outcome as defined in the contract. The eventual signature of the witness depends on whether the exchange was notarized or not.

The notary only sends out receipts, and does not store them; it does not participate in the reputation system process afterwards.

Experience reports, depicted in Figure 3, are submitted to the reputation system by the source about the target, corresponding to the relevant receipts. The identifiers of the actors are as above. Although the main relevance of the notary is in the receipt phase, it is provided in the experience report in order to support credibility analysis: not all of the reputation system participants are required to consider all available notaries trustworthy. When identifying the transaction, the protocol step mentioned in the receipt is omitted here — the single task produces one experience per actor (possibly on both participants in the transaction), but there are multiple receipts that can be used as evidence on a single experience depending on how the transaction progressed. The signature of the source ensures that the experience report cannot be faked: this is required, as submitting a false report that is successfully rebutted has a reputation impact.

- Source actor id, target actor id
- Notary id, if applicable
- Business network model id, eContract id, task id, task counter
- Signature of the source
- Timestamp (time of submission)

**Fig. 3.** The contents of an experience report.

The timestamp here is set by the reputation system when it receives the report for dissemination. It is used to limit the time frame in which rebuttals must be made. Once the time has passed, the source and target are free to dispose of the receipts connected with the experience. In addition, the experi-

ence can then be incorporated into a permanent storage, e.g. transformed into counter increments or similar compound formats that are no longer individually processable.

A rebuttal, depicted in Figure 4, is typically made by the target of the experience to clear its name, as the source has produced the experience to submit. In the case of whitewashing, i.e. undeserved positive experiences, a third party can decide to rebut the experience as well.

- Source id, target id, rebutter id
- Business network model id, eContract id, task id, task counter
- Type of rebuttal: 1) Target rebuttal of false (negative) experience, 2) target noting a failure to report (positive) experience, or 3) target or third party rebuttal to (negative or positive) experience on a nonexistent transaction
- Evidence in form of signed receipts, as applicable
- Signature of the rebutter
- Timestamp (time of rebuttal)

**Fig. 4.** The contents of a rebuttal.

The source and target id refer to their equivalents in the experience, while the rebutter id is either the target's or belongs to a third party whistleblower. The second line identifies the experience being rebutted. The type of rebuttal specifies which kind of response is expected:

For the first type, the target rebutting a (typically negative) experience, the rebutter (i.e. target) must provide any available reciepts proving it has followed the transaction according to contract. If the target's case is sufficient, the source suffers a reputation loss and the experience is removed. If its evidence does not fulfil the requirements of the rebuttal type and protocol, the rebuttal is ineffective. The verification of the evidence can be done centrally by the reputation system; in case the reputation system cannot be sufficiently trusted to follow this protocol fairly, the rebuttals and supporting evidence can be distributed to all the nodes to perform the rebuttal process locally. This choice ties to how information dissemination is organized in the reputation system in general, and is not forced either way by the model of how rebuttals work.

With the second type of rebuttal, the target notifies that the source should have reported a (typically positive) experience on it, but has not. The target, again, must provide the evidence to support its claim. This type forms a special case in defining the appropriate time for the rebuttal: the source should submit the experience without delay after it has provided the target with sufficient evidence, yet we cannot trust a timestamp in the receipt, as it can be set arbitrarily by the source. Instead, the fact that the target is able to present the evidence in the first place indicates that the experience should either be in the system or arriving simultaneously with the rebuttal. If this is not the case and the evidence from the receipts is sufficient, the source suffers negative reputation.

The third type of rebuttal demands a rebuttal response containing proof of the original experience. The target may rebut an experience, typically negative, of a transaction that never happened, or a third-party whistleblower may rebut it, suspecting an undeserved, typically positive experience. In both cases, the burden of proving that the transaction exists and ended as indicated by the experience lies with the source, and if it fails, it suffers a negative reputation impact. The relevant evidence consists of receipts signed by the target or a notary; receipts signed by the source alone are no more credible than the original signed experience report. In the case of the third party rebuttal, it is worth noting that the evidence may itself be a product of collusion between the source and target, in which case the rebuttal may have been appropriate but is still ineffective. The contents of the rebuttal response are depicted in Figure 5.

- Source id, target id, rebutter id
- Business network model id, eContract id, task id, task counter
- Evidence in form of signed receipts

**Fig. 5.** The contents of a rebuttal response.

The rebuttal response can go unsigned, as the signatures of the receipts define its validity. It only needs to be identified as a continuation of the rebuttal before it, and provide the valid evidence. Its arrival time can be compared to the timestamp on the rebuttal, but it has no particular need for a timestamp itself: either it arrives on time to be accepted into the reputation system or it does not. A failure to respond adequately means the original rebuttal is accepted, while a successful response cancels out the rebuttal and the experience remains valid.

## 4 The impact of objective and verifiable experiences

We have defined an objective basis for experiences through associating them with contracts. This gives shared experiences the semantic clarity required to use them in automated decision-making. To ensure the verifiability of experiences, we enforce the fairness of the receipt exchange protocols through the use of notaries. Together, objectivity and verifiability provide a basis for using reputation information to implement social control in the open service ecosystems.

Obreiter discusses different types of nonrepudiable evidence and problems related to them [15]. Two issues in particular must be solved to ensure the acceptability of reputation systems for inter-enterprise collaborations:

- Actors have very limited incentive to provide any evidence of their own misbehaviour, while they may have an incentive to provide unfairly negative reports of their competitors to gain a competitive advantage.
- Two colluding actors can provide an unlimited amount of positive feedback on each other by faking transactions.

We first show how our proposal addresses the first issue by providing the participants appropriate incentives to ensure fair reporting of misbehaviour, and then focus on addressing the second issue through limiting the negative effects of ballot stuffing to other members of the ecosystem. In the third subsection, we compare our work to the state of the art to delineate our contribution.

### 4.1 Ensuring fair reporting on existing transactions

The protocol design for experience reporting and rebuttal aims to reduce the need for rebuttals, limit the incentive for false reporting and omissions by making inaccuracies easier to detect by interested parties, and to balance between giving incentive for third party whistleblowers to rebut likely inaccurate reports, but not to flood the system.

To limit fraudulent positive reports that do not result from collusion, we have chosen the source of a relevant receipt to be the one to submit the related experience to the reputation system. Positive experiences are assumed to be the norm. The target of an experience has motivation to ensure that any deserved positive experiences are reported, possibly also to produce false positive experiences and omit negative experiences. The source, in turn, has motivation to punish the target with honest negative experiences, possibly to report false negatives as well, and omit positive experiences, particularly if the target is a competitor. The target has an interest to report an omission of a positive experience or to clear its name after a false negative experience, and it will know to do this if the reputation system fails to send an expected kind of notification of a reported experience after a transaction. On the other hand, the production of positive experiences about nonexistent transactions with the claimed source requires the source to be around to react to them on time, or a third party whistleblower to take interest; our design sets the source as the reporter to eliminate this issue.

The three types of rebuttals have different motivations and effects. Our goal is to punish false experience reports. To achieve objectivity in this, the exact reputation effects of spreading misinformation are defined in a reputation network contract that must be signed before joining the reputation network. This contract specifies other relevant factors, such as the exact timeframes for rebuttals and their responses, as well. To make the punishment system effective, the rebutters must have an incentive to submit correct rebuttals, and to not submit ungrounded rebuttals.

Actors have a reputation-based incentive to rebut unfairly negative experiences towards themselves. In all cases, the correct or missing experience can be added into the reputation system as a result of a successful rebuttal, in addition to the incorrect experience being removed or marked as invalid. This new experience can be signed by the target, or the centralized reputation system that verifies the evidence, if the latter is available. The target rebutter, in other words, typically gains positive experience as a result, which creates an incentive for it to rebut an experience correctly. Unsuccessful target rebutters can be punished with a negative reputation effect, as they are in a good position to estimate whether they or the source have the evidence to back or counter a rebuttal.

Punishing or awarding third party whistleblowers for their rebuttals is more complicated, however, as their rebuttals at best rely on a guess on whether the experience was a result of a collusion or an attempt to get lucky. The number of third-party rebuttals processed from any actor at a given timeframe can be limited to control the load. A minor reward for a competing service provider is that the artificially inflated reputation of its competitor is reduced, increasing its own reputation in relation to it. A greater incentive can be created by providing a reputation reward for a successful third-party rebuttal, although this in turn must be combined with limitations on the frequency of such rebuttals in order to not create an incentive to flood the reputation system with random rebuttals. In addition, the reputation gains of third-party rebutters should not be as high as the reputation loss of the other actor, as this would create a market for moving reputation from the source to the rebutter. Finally, actors with very bad reputation cannot be reasonably incentivized to behave through threats of further reputation loss, and should therefore be eventually shut out of submitting new information or third-party rebuttals to the reputation system.

Even when a third-party rebuttal is ineffective in the objective sense, disseminating the rebuttal attempt allows reputation system participants trusting the whistleblower more than the source and target together to adjust their local credibility analysis accordingly. These kinds of side effects are an argument for distributing information about the rebuttals to the entire reputation system even if a centralized system could perform the analysis itself.

### 4.2 Addressing collusion to generate positive experiences

As collaborations in the open service ecosystem are impossible to externally observe by third parties unless the collaborating parties allow it, it is possible for two participants to collude to produce positive experiences on each other. To do this, they exchange nonrepudiable receipts according to a protocol, without actually committing concrete resources. They can include an honest notary to observe the exchange and gain further credibility, assuming that the business process does not require costly third-party services to be invoked. Limiting the attractiveness of collusion is a difficult problem. The victim of conducting fake business may also be unobvious: how can positive feedback be harmful?

Let us assume that in a competitive environment, having a higher number of positive experience reports stored on an actor directly influences their probability of being chosen into a collaboration. This, in turn, provides additional opportunities in gaining further positive reputation. The assumption implies that a relative loss of reputation in comparison to one's competitors translates to a monetary loss that is slowly growing over time. Punishment for unfairly causing reputation loss to a victim is intuitively important. When we also consider that within this assumption, a relative gain in reputation in comparison to one's competitors is similar to causing all of them reputation loss, the problem with ballot stuffing in reputation systems becomes acute.

Li et al. propose to solve the issue by assigning a cost to all transactions, which would increase the cost of collusion [11]. For the operational environment

of Pilarcos, however, there is no clear single operator who could collect equal transaction-based fees from all members of the ecosystem, and allowing actors to choose the target of their payments would leave an opening for a more complex collusion that includes a dishonest operator service.

We must therefore resolve this issue within the domain of distributed peer control, and propose to do so by partially breaking the above assumption: positive experiences should not directly improve the probability of being chosen, but positive experiences that are locally found credible would have this effect. Negative experiences should generally have more weight in a decision than positive experiences, and they, in turn, must be backed by evidence.

By valuing local and possibly trusted partners' experiences above random shared experiences, the ecosystem members can limit the gains from collusion between isolated actors; we discuss a selection of different approaches to estimating the credibility of reputation information in earlier work [18, 23]. Local credibility analysis based on e.g. social relations with the information source [4] or the relationship with local experiences [21] is subjective and therefore should only be used to select trustworthy and relevant information sources. It cannot form a basis for second-order punishment, as the reason for a disagreement between two experience sources can be caused by honestly reported discriminatory behaviour.

### 4.3 Comparison to the state of the art

To demarcate our contribution to the state of the art, we compare our solution within Pilarcos to the protocol-based solution proposed by Li et al. [11] on five dimensions: application area, type of third party witness, target of observation, punishment method and implementation requirements. A summary of the comparison is provided in Table 1.

In *application area*, the proposals differ on two levels. Pilarcos operates in a governed open service ecosystem, where transactions are complex and may be long-lived, and varying communication protocols are defined through collaboration contracts. In the proposal of Li et al., the electronic marketplace supports exchanges of goods in simple transactions, and the same protocol suffices for all actors. A more complex environment also means that our view of misbehaviour must be broader, and as a result Li et al.'s theorem on removing actors' incentive to misbehave [11] does not generalize meaningfully into open service ecosystems.

*Third party observation* is implemented in Pilarcos through a notary who acts as a passive witness: it signs receipts, but does not act on them. In contrast, Li et al. have an active arbitrator who is also responsible for judging the outcome and punishing misbehaviour, which gives it more power. Despite this difference, the same impacts [11] and basic limitations of third party protocol-level monitoring apply to both solutions: only protocol-level misbehaviour can be detected through protocol-level monitoring, which means that e.g. compensation processes must be made visible on that level to have a reputation impact.

The *target of observation* in both proposals is the accurate completion of protocols; in Pilarcos they follow contract-defined business processes, while for Li et

| Dimension | Pilarcos | Proposal of Li et al. | Remarks |
|---|---|---|---|
| **Application Area** | Open service ecosystems: collaborators in complex transactions | Electronic multi-agent marketplace: service providers and consumers in simple transactions | Variation in participant roles; additional forms of misbehaviour; the same assumptions do not hold |
| **Third Party Witness** | Notary (passive witness) | Arbitrator (active resolver) | Impact & limitations of 3rd party witness apply to both |
| **Target of Observation** | Accurate completion of different contract-defined business processes | Completion of fixed protocol for purchasing arbitrable & replicatable (intangible) goods | Generalization; some requirements cannot be resolved globally, pushed to contract design instead |
| **Punishment Method** | Distributed: reputation-based peer control, contractual compensation | Centralized: fixed monetary and reputation gains and losses | Different focus and method, more complementary than contradictory |
| **Implementation requirements** | Notaries; protocol design; reputation systems with their own contracts, incl. the rebuttal protocol | Arbitrators; appropriate configurations of service fees collected globally; bank-controlled deposits | Global service fees not feasible; costs of witness protocols are comparable |

**Table 1.** Comparison of Pilarcos to the proposal of Li et al.

al. there is a fixed protocol for purchasing arbitrable and replicatable goods [11]. We therefore consider our solution to be a generalization of their work. Li et al. define arbitrability and replicatability, i.e. that the communication protocol can be repeated for the third-party witness, as requirements for arbitration be effective [11]. We expect that reasonable arbitrability can be reached through business process design for open service ecosystems as well; in some situations, additional third party mediators must be involved to control risks in the collaboration. For particularly trusted partners, these controls can be relaxed.

The *punishment methods* of the systems differ in focus and approach. The reputation-based punishment we propose in this paper is distributed and based on peer control. In addition to reputation-based punishment, Pilarcos contracts contain compensation clauses for misbehaviour, like any business agreements. Li et al. take a centralized point of view both in the arbitrator-based monetary punishment and optionally the broker selecting providers based on their reputation, which is also determined centrally. The approaches seem to be more complementary than contradictory; for example deposits can be applied in high-risk situations to ensure that contractual compensation can be enforced, and reputation-based service selection can be distributed to the actors themselves. We aim for a distributed solution to ensure the viability of the marketplace: as reputation information is worth money, granting a single central actor monopoly over all ecosystem members' reputation is equivalent to creating a new central bank in the ecosystem; this kind of power is disruptive and requires strong control mechanisms to balance for it.

The *implementation requirements* of the two solutions are different in nature. In the proposal of Li et al., the availability of the trusted arbitrator as well as deposit bank and broker is required, and the service fees must be configured globally to minimize incentives to misbehave while maintaining an incentive to

use the system. In addition, all actors are required to make bank-controlled deposits which are held as collateral. The trusted third parties form a single point of failure to the marketplace, which Li et al. aim to distribute more in future work [11].

For Pilarcos, we require the existence of trusted notaries, and push the requirement for designing appropriate arbitrable protocols to specialist business network model designers. In addition, we demand that users of the proposed reputation system, where shared experience information is stored, agree to the reputation network contract. Other costs and impacts of the encompassing Pilarcos system and its trust management system have been discussed in earlier work [17]; we draw additional benefits from the infrastructure for the goal at hand. In contrast to the proposal by Li et al., we estimate that centrally collected global service fees for all transactions are unrealistic in the open service ecosystem, which means that they cannot be generally applied to solve the ballot stuffing problem. Instead, we propose the use of local credibility estimation to reduce the gains from such collusion. This credibility analysis can take into consideration the cost of faking the transaction as well, for example if there is an equivalent of service fees designed into the specific business process.

For implementation cost, the increased messaging for applying third party witnesses to problem situations is not a major cost when optimistic fair exchange protocols are used [13]. The runtime overhead of either system is dominated by cryptographic signing, and we estimate it is not remarkable, considering that the systems involved are capable of running full-blown business protocols.

In summary, our proposed solution is designed for a complex environment; we find that some of the impact of third party witnesses discussed by Li et al. hold for Pilarcos and give implications for collaboration protocol design, while other assumptions made in their game-theoretic analysis [11] cannot hold due to differences in collaboration types. Some aspects of the bank-based punishment proposed by Li et al. are best compared to contractual compensation in the Pilarcos context, and a system for contractually-agreed, deposit-based punishment could well coexist with reputation-based punishment for misbehaviour. Our proposal is distributed and reduces the amount of trust that must be placed on the third party witness or other involved actors.

## 5 Conclusion

We advance a reputation system for inter-enterprise collaborations that is based on objective, verifiable reputation: shared experiences denote whether the collaboration contract was followed or not. To standardize the semantics of experiences in order to make them shareable, we define the reputation effects of different kinds of collaborations in the collaboration contracts. To ensure that false experiences are caught and their submitters punished, an audit trail of the collaboration is produced by signed, nonrepudiable receipts. These receipts can be used to verify whether an experience report is truthful. Objectivity and verifiability go hand in hand: alone, the impact of either remains limited.

The major benefit from this combined approach is the implementation of a two-level sanctioning system, punishing both malicious behaviour and unfair punishments. In the absence of a strong centralized control mechanism, this is necessary to ensure that the service ecosystem does not deteriorate from rampant misbehaviour. In other words, we implement a distributed form of social control in the open service ecosystem.

Trust issues are not entirely solvable by technology alone; in our approach, as well, the final recourse involves lawsuits for contract breaches, and similar infrastructure for ensuring that notaries, i.e. trusted third parties, have an incentive to fulfil their duties. In contrast to the default assumption that trusted third parties are universally trusted, we have strongly limited the amount of trust necessary to place on the proposed notary services.

Objective and verifiable experiences make it possible to punish the spreading of misinformation in the reputation system; due to factors such as collusion to produce positive experiences with little invested effort, they do not remove the need to analyse reputation information locally. Local credibility analysis of all incoming experiences remains a central technical recourse against misinformation: it is not necessary nor prudent to accept all experiences as equal, be they subjective or objective.

## References


1. Blaze, M., Feigenbaum, J., Keromytis, A.D.: KeyNote: Trust management for public-key infrastructures (position paper). In: Proceedings of Security Protocols: 6th International Workshop, Cambridge, UK, April 1998. pp. 59–63. Springer-Verlag, LNCS 1550/1998 (Apr 1998)
2. Dondio, P., Longo, L., Barrett, S.: A translation mechanism for recommendations. In: Trust Management II. IFIP International Federation for Information Processing, vol. 263, pp. 87–102. Springer, Pisa, Italy (May 2008)
3. Fehr, E., Fischbacher, U.: The nature of human altruism. Nature 425 (Oct 2003)
4. Gal-Oz, N., Gudes, E., Hendler, D.: A robust and knot-aware trust-based reputation model. In: Trust Management II. IFIP International Federation for Information Processing, vol. 263, pp. 167–182. Springer, Pisa, Italy (May 2008)
5. Hasan, O., Brunie, L., Pierson, J.M., Bertino, E.: Elimination of subjectivity from trust recommendation. In: The 3rd IFIP International Conference on Trust Management (TM 2009). pp. 65–80. West Lafayette, IN, USA (2009)
6. Kaur, P., Ruohomaa, S., Kutvonen, L.: User interface for trust decision making in inter-enterprise collaborations. In: Proceedings of the Fifth International Conference on Advances in Computer-Human Interactions (ACHI 2012). pp. 122–127. IARIA, Valencia, Spain (Jan 2012)
7. Kutvonen, L., Metso, J., Ruohomaa, S.: From trading to eCommunity management: Responding to social and contractual challenges. Information Systems Frontiers (ISF) - Special Issue on Enterprise Services Computing: Evolution and Challenges 9(2–3), 181–194 (Jul 2007)
8. Kutvonen, L., Ruokolainen, T., Metso, J.: Interoperability middleware for federated business services in web-Pilarcos. International Journal of Enterprise Information Systems, Special issue on Interoperability of Enterprise Systems and Applications 3(1), 1–21 (Jan 2007)

9. Kutvonen, L., Ruokolainen, T., Ruohomaa, S., Metso, J.: Service-oriented middleware for managing inter-enterprise collaborations. In: Global Implications of Modern Enterprise Information Systems: Technologies and Applications. pp. 209–241. Advances in Enterprise Information Systems (AEIS), IGI Global (Dec 2008)
10. Lee, S., Sherwood, R., Bhattacharjee, B.: Cooperative peer groups in NICE. In: Twenty-Second Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2003). vol. 2, pp. 1272–1282. IEEE (Apr 2003)
11. Li, Q., Martin, K.M., Zhang, J.: Design of a multiagent-based e-marketplace to secure service trading on the Internet. In: Proceedings of the 13th International Conference on Electronic Commerce. Liverpool, UK (Aug 2011)
12. Mehandiev, N., Grefen, P. (eds.): Dynamic Business Process Formation for Instant Virtual Enterprises. Advanced Information and Knowledge Processing, Springer (Jun 2010)
13. Micali, S.: Simple and fast optimistic protocols for fair electronic exchange. In: Proceedings of the twenty-second annual symposium on Principles of distributed computing (PODC '03). pp. 12–19. ACM (2003)
14. OASIS Web Service Secure Exchange TC: WS-Trust 1.3 OASIS Standard. OASIS (Mar 2007)
15. Obreiter, P.: A case for evidence-aware distributed reputation systems overcoming the limitations of plausibility considerations. In: Proceedings of Trust Management: Second International Conference, iTrust 2004, Oxford, UK, March 29–April 1, 2004. pp. 33–47. Springer-Verlag, LNCS 2995/2004 (Mar 2004)
16. Resnick, P., Zeckhauser, R.: Trust among strangers in internet transactions: Empirical analysis of eBay's reputation system. In: The Economics of the Internet and E-Commerce. Advances in Applied Microeconomics, vol. 11, pp. 127–157. Elsevier Science, Amsterdam (2002)
17. Ruohomaa, S., Kutvonen, L.: Trust and distrust in adaptive inter-enterprise collaboration management. Journal of Theoretical and Applied Electronic Commerce Research 5(2), 118–136 (Aug 2010)
18. Ruohomaa, S., Kutvonen, L., Koutrouli, E.: Reputation management survey. In: Proceedings of the 2nd International Conference on Availability, Reliability and Security (ARES 2007). pp. 103–111. IEEE Computer Society, Vienna, Austria (Apr 2007)
19. Ruokolainen, T., Ruohomaa, S., Kutvonen, L.: Solving service ecosystem governance. In: Proceedings of the 15th IEEE International EDOC Conference Workshops. pp. 18–25. IEEE Computer Society, Helsinki, Finland (Aug 2011)
20. Srivatsa, M., Xiong, L., Liu, L.: TrustGuard: countering vulnerabilities in reputation management for decentralized overlay networks. In: WWW '05: Proceedings of the 14th International Conference on the World Wide Web. pp. 422–431. ACM Press, New York, USA (May 2005)
21. Teacy, W.T.L., Patel, J., Jennings, N.R., Luck, M.: TRAVOS: Trust and reputation in the context of inaccurate reputation sources. Autonomous Agents and Multiagent Systems 12(2), 183–198 (Mar 2006)
22. Wilson, M.D., Arenas, A., Schubert, L., et al.: Trustcom framework V2, Deliverable D29, D35, D36. Tech. rep., TrustCoM WP27 (Jan 2006)
23. Yao, Y., Ruohomaa, S., Xu, F.: Addressing common vulnerabilities of reputation systems for electronic commerce. Journal of Theoretical and Applied Electronic Commerce Research 7, 1–15 (Apr 2012)