Journal of Theoretical and Applied Electronic Commerce Research
ISSN 0718–1876 Electronic Version
VOL 5 / ISSUE 2 / AUGUST 2010 / 97-117

This paper is available online at
www.jtaer.com
DOI: 10.4067/S0718-18762010000200007

The State of the Art in Trust and Reputation Systems: A Framework for Comparison

Zeinab Noorian[1] and Mihaela Ulieru[2]

University of New Brunswick, Faculty of Computer Science
[1] z.noorian@unb.ca, [2] ulieru@unb.ca

Received 15 February 2010; received in revised form 16 June 2010; accepted 18 June 2010

Abstract

We introduce a multidimensional framework for classifying and comparing trust and reputation (T&R) systems. The framework dimensions encompass both hard and soft features of such systems including different witness location approaches, various reputation calculation engines, variety of information sources and rating systems which are categorised as hard features, and also basic reputation measurement parameters, context diversity checking, reliability and honesty assessment and adaptability which are referred to as soft features. Specifically, the framework dimensions answer questions related to major characteristics of T&R systems including those parameters from the real world that should be imitated in a virtual environment. The proposed framework can serve as a basis to understand the current state of the art in the area of computational trust and reputation and also help in designing suitable control mechanisms for online communities. In addition, we have provided a critical analysis of some of the existing techniques in the literature compared within the context of the proposed framework dimensions.

Key words: Computational Trust, Reputation Formalization, Reliability Assessment, Online Communities, T&R Taxonomy

Journal of Theoretical and Applied Electronic Commerce Research
ISSN 0718–1876 Electronic Version
VOL 5 / ISSUE 2 / AUGUST 2010 / 97-117

This paper is available online at
www.jtaer.com
DOI: 10.4067/S0718-18762010000200007

1 Introduction

Overcoming the inherent uncertainties and risks of the open electronic marketplace and online collaboration systems requires the establishment of mutual trust between service providers and service consumers. In fact, one of the main concerns of such environments is how the systems' resistance against self-interested participants can be enhanced and in what way their actual deceitful intentions can be understood and revealed. To address these concerns, Trust and Reputation (T&R) systems are developed to evaluate the reliability and credibility of the participants such that recommendation can be made when needed. Generally stated, the underlying goal of all T&R systems is to predict the trustworthiness and proficiency of peers in future actions based on the information gathered from their past behavior in the environment and their peers' view towards their history [22]. Trust can be deduced from both *individual* and *social* perspectives [5]-[6]. Individual trust is due to direct experiences of transaction partners while social trust is calculated from third-parties experiences, which might include both honest and misleading opinions. T&R systems provide individuals with tools and techniques to deliberately solicit reputation information from peers in order to construct reasonable models of reputation for each participant.

Obviously, it is unrealistic to assume that first-hand opinions (individual trust information) are ubiquitously available and that all the recommendations received from witnesses are truthful. Therefore, this shows why T&R systems are important in an open and dynamic environment. To clarify, any T&R system should be equipped with the ability to assess the behavior of the community's participants and detect deceitful behavior of those who have tendencies to gain benefits by conducting dishonest activities [29], [33]-[34]. Besides, these systems should observe the trends and behavioral models of particular service providers in order to identify discriminating attitudes and fraudulent activities [9]. Furthermore, in order to accurately predict the reputation of peers in particular service provisioning scenarios, T&R systems might execute context diversity checking to determine the similarity rate of potential service-level agreement with the previously negotiated contract and thus measure the influence degree of recommendations in the trustworthiness computation process [2].

In this paper, we introduce a framework for classifying and comparing trust and reputation systems. Specifically, the framework dimensions answer important questions related to the major characteristics of T&R systems including those parameters from the real world that should be mirrored in a virtual environment in order to support trust in a virtual community. For instance, the virtual community might need to inherit the dynamicity and fuzziness qualities of the real world to evaluate reliability and efficiency in a virtual environment. Our work differs from others [1], [9], [20], [22] in several ways. First we take a coherent approach to thoroughly study current literature in trust and reputation systems. Second, we provide a critical analysis of the performance of these systems with respect to the features presented in the framework. We believe that the proposed framework can be used as foundation for advancing the research agenda in T&R systems.

We begin with an extensive overview of six well-known trust and reputation systems. Subsequently, we provide the detailed description of the framework and its respective dimensions. Afterwards, we thoroughly compare the existing T&R systems based on the proposed framework and examine their suitability on different environments and application domains. We also analyze their pros and cons by effectively addressing some advanced features of the framework. Finally, we conclude the paper by explaining some of the open problems in this field.

2 Trust and Reputation Systems

Given that T&R systems are context sensitive, the design of different existing models and systems has been dependent on the target domain and the related specific requirements. In the following, we review some of the available systems and discuss how they are able to fulfill their goals. We attempt to select the T&R systems with different approaches and techniques in dealing with the intrinsic challenges of the open environment. More explicitly, the chosen T&R systems have distinguished features in dealing with inherent dynamicity of the open environment, evaluating the honesty and reliability of participants, and calculating the reputation score. Such diversity enables readers to obtain decent understanding about existing literatures in trust and reputation systems and observe their applicability in virtual community.

2.1 FIRE Model

In the FIRE model [6]-[7], trust is evaluated within the context of a different number of information components: 1) *Interaction Trust* (IT) that is built from the direct self experience of an agent with the other agents; 2) *Witness Reputation* (WR) that is based on the direct observation of an agent's behavior by some third-party agent; 3) *Certified Reputation* (CR), being one of the novelties in the FIRE model, consists of certified references disclosed by third-party agents. Such information is made available upon request of an inquiring agent. The CR component is desirable in the absence of direct interaction and when witnesses are self-interested and reluctant to share their experiences. Moreover, the use of CR enables agents to be freed from the cost of locating witnesses while their confidence rate of the anticipated trust value in not compromised. 4), the last component is *Role-based Trust* (RT), which models the trust across predefined role-based relationships between two agents, e.g., (owned by the same

This paper is available online at
www.jtaer.com
DOI: 10.4067/S0718-18762010000200007

company, friendship relationship, team-mate relationship) [7]. In this case, by defining and updating these roles in open Multi-Agent Systems (MAS) as well as assigning the expected trust value and belief strength (of relying agent) on them, RT is able to contribute in trustworthiness prediction for future interactions. It is worthwhile to mention that the significance of each component in the composite trust formula is adjusted automatically according to unforeseen changes in the environment. In this trust model, each component owns a trust formula with relevant rating weight function to determine the quality of ratings tailored to its responsibility. For instance, it seems sufficient for IT to design the weight function solely based on the recency of ratings whereas WR and CR have to take the credibility of rating into account as well. To address this requisite, FIRE has developed a mechanism to filter out the inaccurate reports revealed by unfaithful witnesses and penalises them accordingly. In so doing, it defines an *inaccuracy tolerance threshold* (L) to specify the maximal permitted differences between the actual performance and witness rating. Credibility of each rating is tuned to be inversely proportional to the differences, i.e., the higher the differences are, the lower the credibility [5]. Furthermore, the FIRE model defines a reliability measure to calculate the confidence level of an agent in believing that another agent can perform as expected. In general, it provides two types of reliability: *rating reliability*, which depends on the number of available ratings with high values, which depict the expected performance of the target agent. The other type is *deviation reliability,* which intends to examine the volatility of the target agent in accomplishing an agreement. Basically, it calculates the deviation of ratings around the produced expected value [5]. Intuitively, if the target agent showed an inconsistent behavior while countering a different requesting agent, its reliability value will be gradually affected negatively.

Note that the FIRE model inherits IT and multiple-criterion rating systems from the REGRET [26] reputation system and for the purpose of seeking and locating the relevant witnesses in WR, it is inspired by the decentralized approach of Singh and Yu's referral network [34] and implements a variant of their system.

2.2 REGRET

REGRET [24], [26] is a decentralized trust and reputation model designed for complex e-commerce environments where various types of agents with different social relationships play important roles. With the help of a social structure called *sociogram,* it is able to model the social relationships such as cooperation, competition and trade in a graph where the nodes represent the participants and the edges denote the nature of their relationship. This T&R system is based on a three-dimensional reputation model: 1) *Individual dimension* or subjective reputation which calculates trust based on the direct impressions of an agent received from Service provider (SP) and prioritizes its direct experiences according to their recency; 2) *social dimension* which is designed to estimate the trustworthiness of SP in case the direct experiences are insufficient or the agent has newly joined the environment. This dimension is itself divided into three specialized types of reputation depending on the information sources. First, *witness reputation* which calculates reputation based on the information coming from the witnesses adjacent to this agent. Here, adjacency is defined as an indication that some form of relationship between two agents exists. Second, *neighbourhood reputation* that measures the reputation of individuals who are neighbours with the agent being evaluated by considering their social relationships and third, *system reputation* which assesses the trustworthiness of SP based on the general role that it plays in the sociogram. In order for REGRET to be able to calculate social reputation, it must first identify appropriate witnesses in the e-commerce environment. For doing so, it applies graph theory techniques to the sociogram to locate the most appropriate witnesses and examines their social relationships with the agent being evaluated. Furthermore, by presenting the social relationship in the form of fuzzy rules, REGRET is able to determine the honesty and credibility of the reported observations thus assigning suitable weights to them. For instance, it may declare that (IF *the competition relation of witness A with the target agent is very high*, THEN *its recommended reputation value should be very bad*). 3) The third reputation dimension of REGRET is the *ontological dimension,* which adds the possibility of combining different aspects of reputation to calculate a complex one [5].Note that in the last two dimensions, the agent recorded impressions are linked to single behavioral aspects and do not provide general ratings. However, with the help of the ontological structure, each agent is capable of determining the overall reputation of a particular SP by assigning the appropriate influence degree to each aspect tailored to its demand. In addition to the reputation value, REGRET comes with a reliability measurement which reflects the confidence level of the produced reputation value. Similar to SPORAS [36] and FIRE [6], reliability measurement is calculated from a combination of two factors: the number of available impressions and the variability of the impression values. In order to boost the accuracy of the reliability measure, REGRET defines the intimacy level of interaction which indicates the maximum number of impressions required for a close relationship. As the number of impressions grows, the reliability degree increases until it reaches a certain intimate value. Afterwards, reliability is not affected by the increment of the intimate parameter. It is important to mention that the value of the intimate parameter is dynamically adjustable depending on the interaction frequency of individuals as well as the quality of impressions [20].

2.3 T&R Model by Yu and Singh

The T&R model designed by Yu and Singh [20], [35] contains various distinctive features which surpass other available models in some contexts. This model of reputation management exploits two information components. The first one contains the agent's local belief built as a result of its direct interaction with other agents. The second one includes the testimonies of third-parties that can be beneficial in the absence of local ratings. In this model, in order

This paper is available online at
www.jtaer.com
DOI: 10.4067/S0718-18762010000200007

to estimate the total belief regarding the trustworthiness of a particular agent, the requesting agent combines a local belief in conjunction with third-party testimonies to achieve a more accurate evaluation.

Furthermore, Yu and Singh propose a novel trust network which intends to locate the most appropriate witnesses in a multiagent system. In this model, each agent is surrounded by a number of acquaintances among whose subsets can be neighbours. When a requesting agent wants to evaluate the trustworthiness of a particular agent, it will send a query to the neighbours of that agent asking for their perception regarding the target agent. Unless the neighbours have not had any direct experiences with that agent they respond by their testimonies; otherwise, they will reply by returning a series of referrals. The number of referrals is limited by the *branching factor* and *depthLimit* parameters [34] so as to limit the effort expended in pursuing referrals. This process successfully terminates if an adequate number of ratings are received and it encounters failures when the *depthLimit* is reached and neither ratings nor referrals are gathered [35]. Note that each individual agent maintains a two-dimensional model of each acquaintance. The first dimension indicates their ability to act in a trustworthy manner, which is called *expertise* and the other one signifies their *sociability* in referring to suitable trustworthy agents. Depending on their competency in fulfilling either of the above-mentioned qualities, acquaintance models are modified to reflect their actual performance to be used in future interactions.

The other major concern of this model is dealing with deceptive agents who deliberately disseminate misinformation through network for their self-interest. The proposed model considers three types of deceptions [34]: *complementary*, *exaggerative positive* and *exaggerative negative*. This classification is based on the behavioral model of the participants in giving ratings. For instance, if agents intentionally give controversial ratings, they may be detected as malicious agents with complementary model of deception. Such agents will lose credibility in the update phase. Similarly, an agent with exaggerative positive tendency acts rather untruthfully in the system. To clarify, even if it is not fully satisfied with the performance of a particular agent, it provides a higher rating than it actually experienced. The possible motivation for this behavior could be receiving of a commission from the other agent. Consequently, the credibility of this agent is reduced in proportion with its dishonesty. Moreover, depending on the system's circumstances, this model defines an *exaggeration coefficient,* which determines how much agents could lie before they are considered as being exaggerative and not a complementary deceptive agent. Note that after the actual interaction with the recommended target agent, the requesting agent re-calculates the weight of the witnesses and updates their credibility degree for subsequent reputation prediction processes.

Finally, in order to tackle with the uncertainty factors inherent in open MAS, this reputation management model benefits from the Dempster-Shafer theory of evidence [3] as an underlying computational framework. According to this theory, lack of belief does not necessarily imply disbelief in the system. Thus, instead of assuming total disbelief as initial value for newcomers, it is replaced by a state of uncertainty. In other words, with the help of the theory of evidence, Yu and Singh's model is able to differentiate between having a bad reputation and no reputation at all [35], [27]. Moreover, to predict total belief it utilizes Dempster's rule of combination [3] as an aggregation method which combines evidences to compute new a belief value. In addition, this model describes a variant of the Weighted Majority Algorithm (WMA) [15] in order to fine tune the weight of advisers for the purpose of deception detection after actual successful or unsuccessful interactions.

2.4 TRAVOS

The TRAOVS (Trust and Reputation model for Agent-based Virtual Organizations) system is developed to ensure high-quality interaction between the participants of a large open system [19]. It exploits two information sources to assess the trustworthiness of the participants: *Direct Interaction* and *Witness Observation*. To derive trust, this model relies greatly on its direct experiences and refuses to combine others' opinions unless they are really required. For this purpose, it provides a *confidence* metric to determine whether the personal experiences are sufficient to make an acceptable judgment with respect to a particular SP or not. If not, it disseminates queries to obtain additional observations from other witnesses who claim to have had previous interaction with that certain SP.

Specifically, this T&R model utilizes a single rating system such that the outcomes of the interactions are summarized in a single variable which indicates an overall performance. Here, witnesses share the history of their interactions in a tuple which contains the frequency of successful and unsuccessful interaction results.

Moreover, in order to deal with inaccurate reputation providers, TRAVOS takes advantage of an exogenous approach presented in [9], [31]. According to this approach, instead of calculating the reliability of the provided recommendation based on its deviation from mainstream opinions, it calculates the probability that a particular correspondent provides accurate reports given its past opinions and proportionally adjusts the influence of its current observations afterwards. To clarify, as a first step, TRAVOS considers the actual results of all previous interactions with collection of SPs in which the agent provided similar observations. Then, by means of comparing the variables of their beta distributions it is able to measure the degree of accuracy of that certain agent. That is, truster agent constructs a beta distribution of the rater's current opinion and calculates the relevant expected value E^r. It also builds the beta distribution of all the previous outcomes in which the rater has provided similar opinions and estimates its expected value E^0 as well. Then, by means of comparing their corresponding expected values, TRAVOS is able to conclude the honesty and accuracy of a rater's current observation (c.f. [19], [29]). In the second step, this T&R system attempts to decrease the effect of unreliable opinions on a final computed reputation value. An

This paper is available online at
www.jtaer.com
DOI: 10.4067/S0718-18762010000200007

untruthful agent could considerably affect the reputation of the queried SP by providing a huge number of unfair ratings. This problem arises because of its method of reputation combination, which is based on a simple summation of all the provided opinions. To rectify this, TRAVOS adopts techniques to reduce the amount of ratings unless the accuracy degree of the opinion provider is very high.

2.5 PeerTrust

PeerTrust [22], [32]-[33] is a coherent dynamic trust model with unique characteristics tailored for peer-to-peer e-commerce communities. For advanced assessment and quantification of peer's trust value in constantly evolving environments, this model customises a variety of common factors: 1) *feedback* which is a judgment of other peers regarding target peer; 2) *feedback scope* such as the amount of transactions the peer experienced with others; 3) *credibility factor* for evaluating the honesty of feedback sources 4*) transaction context factor* such as time and size of transactions which could act as defense mechanism against delicate fraudulent activities; and 5) *community context factor* that addresses the feedback incentive problem. This model proposes an innovative composite trust metric that incorporates the described parameters to enhance accuracy and reliability of predicted trustworthiness.

One of common way for malicious participants to undetectably continue sabotaging in the system is maintaining their general trust value at a certain level by increasing the transaction volume which hides the effect of their frequent frauds. To alleviate the effect of those malicious attacks resulted from increase in transaction volumes it combines the first two parameters such that instead of simply aggregating generic feedback values, it equips witnessed-peers with the ability to disseminate their degree of satisfaction by calculating the average amount of successful outcomes that they experienced.

Besides, to ensure the quality of the reputation information, peers are equipped with credibility measures to calculate the credible amount of satisfaction. In doing so, PeerTrust defines the personalised similarity measures [32]-[33] which compute feedback similarity rate between the evaluating peer and opinion providers over a common set of peers with whom they have had previous interaction. Since trustworthy peers consistently act honestly as a role of feedback provider and do not become affected by malicious intentions such as jealousy and negative competitive attitude, in addition, this model also advocates that the trust metric can be alternatively served as a credibility measure under certain circumstances.
Evidently, one of the significant parameters which is widely neglected in T&R systems is transaction context. More explicitly, PeerTrust emphasizes that the aggregation of feedback which are only based on the credibility of their correspondents cannot efficiently reflect the trustworthiness of the agents. Thus, it incorporates various aspects of transaction such as its size, time and category under Transaction Context factors to model participants' intentions and potential fraudulent activities in the trustworthiness measurement.

Furthermore, it is widely agreed that feedback are one of the foundations of T&R systems such that these systems cannot perform effectively unless they have access to a sufficient amount of feedback [14]. Therefore, to stimulate participants' cooperation, PeerTrust embeds a reward function, called the community context factor, into the trust metric to encourage peers to persistently provide votes about others' performance.

The dynamic and distributed nature of peer-to-peer systems necessitates an optimized and adaptive design of the peer location approach. To operationalize this goal, this model provides each peer with a trust manager and a data locator engine which are responsible for feedback submission and retrieval aside from trust evaluation over the underlying network.

2.6 BRS

Jøsang et al. [8], [11] have proposed the flexible and adaptive Bayesian Reputation System(BRS) which supports both binomial and multinomial rating models to allow rating provision happen in different levels of precision well-suited for open dynamic environment. Theoretically, multinomial BRS is based on computing reputation scores by statistically updating the Dirichlet Probability Density Function (PDF) [13]. More explicitly, in this context, agents are allowed to rate other peers within any level from a set of predefined ratings levels. In contrast, in binomial BRS which is based on Beta distribution, the agents can only provide binary ratings for the others. That is, in multinomial BRS the reputation scores do not solely reflect the general quality of service; but are also able to distinguish between the case of polarized ratings and the case of average ratings [12]. Evidently, such differences are not noticeable in binomial ratings, resulting in uncertainty and low confidence rate in aggregated reputation score and also might prohibit the reputation scores to converge to specific values [30]. Furthermore, multinomial BRS allows the input ratings to be provided based on both discrete and continuous measures to reflect a rater's opinion more accurately when required. To operationalize this goal, it exploits the fuzzy set membership functions to transform continuous ratings into discrete ones in order to provide compatible inputs for BRS [12]. Both systems use the same principle to compute the expected reputation scores, namely by combining previous interaction records with new ratings.

Moreover, BRS appears to be promising method to foster trust amongst strangers in an online environment. It takes an innovative approach which enables trustee agents to evaluate the sincerity of the ratings provided by recommendation agents outside of its control. As such, it uses the endogenous discounting method to exclude such

Journal of Theoretical and Applied Electronic Commerce Research
ISSN 0718–1876 Electronic Version
VOL 5 / ISSUE 2 / AUGUST 2010 / 97-117

This paper is available online at
www.jtaer.com
DOI: 10.4067/S0718-18762010000200007

advisers whose probability distribution of ratings significantly deviate from the overall reputation scores of the target agent [31]. That is, it dynamically determines upper and lower bound thresholds in order to adjust the iterated filtering algorithm's sensitivity tailored to different environmental circumstances. For instance, if the majority of participants act deceitfully in the environment, the lower bound would be set to a higher value so as to increase the sensitivity of the BRS which can lead to the exclusion of more unfair raters. Besides, in order to deal with dynamicity in the participant's behavior, BRS provides a *longevity* factor which determines the expiry time of the old ratings and gives greater weight to more recent ones. As such, it defines a recursive updating algorithm based on the longevity factor to update the participants' reputation scores in certain time intervals. It is noteworthy to mention that, this recursive algorithm also provides a measure to calculate convergence values for the reputation scores [11].

Table 1: Summary of the selected T&R systems

T&R System's Name	References	Distinguishing Features
FIRE	(T.D Huynh, N.R. Jennings, N.R.Shadbolt, 2006)[6]	Designed for Multiagent system, exploits four information sources, handles the bootstapping problem of newcomers, filters out inaccurate reputation information, attempts to differentiate between dishonest and mistaken agents, provides compound reliability measures, employs a multi-criterion rating system ,supports dynamism in open MAS.
REGRET	(Jordi Sabater and Carles Sierra, 2002)[25]	Designed for complex e-commerce systems, develops *sociogram* to model social relationships, supports neighbourhood & system reputation, and provides ontological dimensions to combine various behavioral aspects of reputation. Evaluates witness honesty through fuzzy rules. Provides reliability measure; employs a multi-criteria rating system.
Model by Yu& Singh	(B. Yu, M.P. Singh,2003)[34]	Suitable for MAS, proposes novel trust & referral network, detects three models of deceptions. Provides credibility measures pertaining to each model. Differentiate between agents having bad reputation or no reputation using Dempster-Shafter theory of evidence. Supports dynamism in open MAS.
TRAVOS	(W. T. L. Teacy, J. Patel, *et al,*2006)[29]	Designed for large-scale open system, provides two information sources, exploits a probabilistic approach to determine credibility of witnesses, provides confidence metric and reliability measure for direct interaction information sources; Employs a single-rating system.
PeerTrust	(L. Xiong and L. Liu,2004)[33]	Designed for P2P e-commerce systems, provides two methods as credibility measures, supports transaction context and community context factors in trust metric, and employs an adaptive architecture for peer location. Supports dynamism in peer2peer systems. Attempts to address bootstrapping problem. Support a single-rating system.
BRS	(A.Jøsang, , 2002) [8]	Suitable for open dynamic environment, support binomial and multinomial ratings models, address bootstrapping problem by considering the quality of community in the marketplace, provide iterated filtering algorithm which can effectively reveal deceptive intentions if the majority of participants act honestly, utilize longevity factor to discount ratings as time progress, enable participants as buyers and sellers to adaptively change their behavior in order to increase their own benefits.

3 The Proposed Comparison Framework and its Dimensions

The proposed comparison framework targets the distinctive aspects of T&R systems mapping them against a variety of features along several dimensions, (Figure 1). Generally, the dimensions are classified in two main categories: *soft features* and *hard features*. Such attributes and traits which help to enhance the performance of the system and quality of outcomes are defined within the context of soft features. Moreover, the notion of robustness in the open communities is well-addressed through specific dimensions of soft features. That is, certain dimensions of soft features are introduced in response to potential attacks that threaten the open dynamic environment. For instance, the *reliability and honesty assessment* dimension implies potential vulnerabilities such as collusion, value imbalance, discrimination and playbook [10], [16] in an open dynamic environment. On the other hand, hard features that encompass the essential engines and possible architectures of T&R systems represents the underlying characteristics which are fundamental to establish any T&R systems.

Journal of Theoretical and Applied Electronic Commerce Research
ISSN 0718–1876 Electronic Version
VOL 5 / ISSUE 2 / AUGUST 2010 / 97-117

This paper is available online at
www.jtaer.com
DOI: 10.4067/S0718-18762010000200007

In the following, we provide an extensive description of each dimension of the framework.

3.1 Hard Features

The hard features of T&R systems can be categorized along the following dimensions:

3.1.1 Rating Approaches

One of the underpinning features of any T&R system is the ability to qualify the performance of transaction partners immediately after actual interactions. From this perspective T&R systems rank the quality of delivered services according to the agreed contracts as follows:
Single-criterion (binary) rating system (R1a):In single-criterion rating systems which are also referred to as binary rating systems, participants reveal their general opinion with regards to a target subject. In this approach, requesting agents are asked to publish their subjective view of the overall performance of their peers. Since the interaction aspects of transaction partners are hidden, even if the recommendation agents are honest and trustworthy the exposed information is not very reliable and accurate.

Multiple-criterion rating system (R1b): Unlike binary rating systems, here agents tend to release different aspects of their interactions along with the corresponding evaluation values. It qualifies requesting agent to learn about the agreed criteria of former transaction partners and to compare them with its requirements in order to make informed and accurate decisions. For instance, instead of inquiring about the performance aspect only, requesting agent receives various communicating aspects such as price, delivery and quality of service. Afterwards, it prioritizes them to predict the expected interaction result more precisely.

3.1.2 Witness Locating Approaches

In order to locate the recommendation agents and truthful service provider, every T&R systems should develop well-defined and expressive mechanisms which provide the participants with the ability to identify reliable and proficient witnesses and share contextual information of their previous negotiated contacts regarding certain transaction partners. We emphasize three approaches for this feature:
Centralized (R2a):In centralized reputation systems, information about the QoS of a particular SP is stored in a central database. For this reason, after every transaction agents are asked to report the performance of their transaction partners to a central authority. The accumulated ratings are evaluated using an employed reputation calculation engine to derive the expected reputation score for every participant and then make them publicly available [9]. This approach is widely used in online auction sites like eBay (Site 1) and Amazon (Site 2) where the buyers learn about the trustfulness of sellers before initiating direct interaction. However, this approach suffers from drawbacks inherent in its nature. Centralized reputation systems is in complete contradiction with the characteristics of dynamic environment where the population of the participants vary over time and subsequently when the number of agents grows in the distributed environment, the cost of reporting reputation scores to the central authority becomes enormous. Moreover, the centralized system is practical when it is approved and trusted ubiquitously by all participants; however, this assumption is quite optimistic since there is no ultimate authority for all agents. Furthermore, as described before, the verification of the quality of received feedback is not a trivial task. This problem gets worse in centralized reputation systems since it never directly involves any transaction; thus, malicious agents can simply fool the system by providing deceitful ratings [23].

Decentralized (R2b): In a decentralized reputation system there is no central location for submitting the transaction feedback and for querying the expected reputation score of a particular destination agent [9]. Instead, each agent is capable of storing the ratings locally and manipulates them when required. Moreover, in this approach participants are equipped with the reputation computation engines tailored to their demands so as to calculate and validate the trustfulness of their transaction partners. Examples of the environment types that are well-suited to accommodate distributed reputation systems are peer-to-peer networks like Freenet (Site 3) - and open MAS. In such evolving environments, participants can freely seek for the right advisers and verify the quality of their reputation reports in order to make sensible decisions. Note that unlike the centralized approach, in this system there is no single point of failure which assures the accessibility of services in all situations.

Hybrid (R2c): Such reputation systems attempt to inherit both the benefits of centralized and distributed architectures. An example of trust model which follows this approach was developed by Radu Jurca [14]. To overcome the single point of failure problem, hybrid reputation systems employ sets of brokers responsible for trading reputation information upon request. Their services can be offered in a variety of qualities. Therefore, requesting agents should be given techniques for distinguishing between honest and fraudulent brokers. However, even though the hybrid reputation system is able to perform considerably better in dynamic environments compared with the centralized approach, still it cannot deal with the scalability problem of such environments. For instance, in case the population of the participants grows very large, this system incurs tremendous costs to find and locate appropriate brokers. Thus, this shortcoming deprives it from being widely used as an underlying framework in T&R systems.

Journal of Theoretical and Applied Electronic Commerce Research
ISSN 0718–1876 Electronic Version
VOL 5 / ISSUE 2 / AUGUST 2010 / 97-117

This paper is available online at
www.jtaer.com
DOI: 10.4067/S0718-18762010000200007

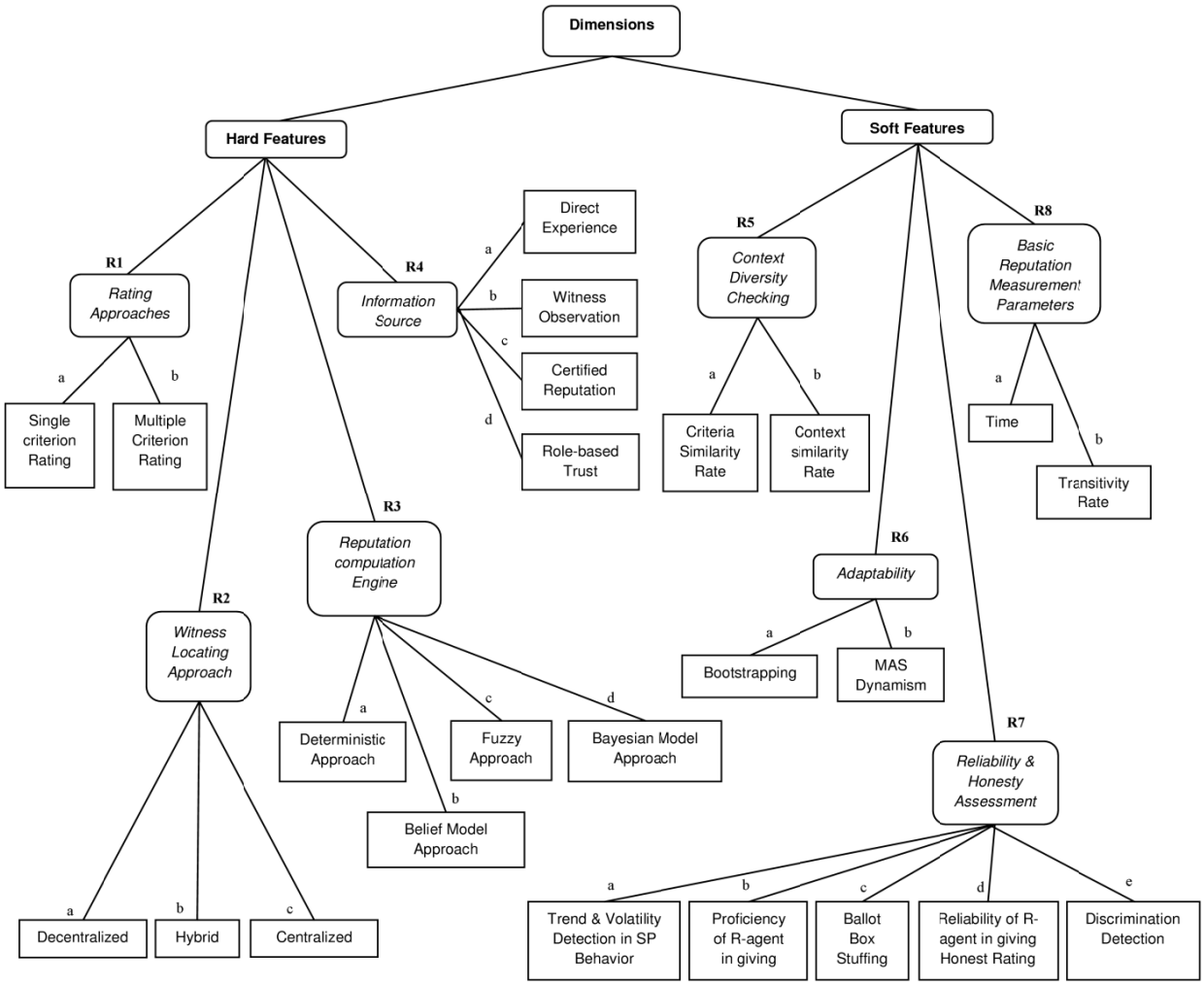

Figure 1: The comparison framework and its dimensions

3.1.3 Reputation Computation Engines

In T&R systems, participants should be able to evaluate the reputation of the potential service provider before actual interaction. Thus, such systems may exploit suitable *reputation computation engines* to aggregate variety of soft feature's values as well as the revealed ratings to predict the trustworthiness of certain transaction partners.
In following the most well-known approaches are presented.

Deterministic approach (R3a):In the deterministic approach, trust values are calculated from handcrafted formula to yield the desired results. In fact, the flexibility of this approach enables T&R systems to define a composite trust metric to aggregate the essential parameters and factors they have been considered in their models. Basically, they include credibility of witnesses and time or recency factor as the main variables of the systems. However, some advanced models may take other important variables such as the transitivity rate and context & criteria similarity rate into account as well [6], [33]. For instance, [2] proposes an aggregation method considering the following parameters:

- Witness Trustworthiness Value (WTV): $(TC_A(i,n,c)$, $t = (1,2,\ldots n)$
- The reputation information: $R_A(i,k,n,c)$
- The time weighting factor of the provided opinion : $TF_A(i,n,c)$
- Transitivity rate α_i ,

Journal of Theoretical and Applied Electronic Commerce Research
ISSN 0718–1876 Electronic Version
VOL 5 / ISSUE 2 / AUGUST 2010 / 97-117

This paper is available online at
www.jtaer.com
DOI: 10.4067/S0718-18762010000200007

And the trust metric is presented as follows:

$$T_{BA}(K,n,c) = \alpha_1 \sum_{i}^{L} (TC_A(i,n,c) * R_A(i,k,n,c) * TF_A(i,n,c)$$
$$+ \alpha_2 \sum_{j}^{L} (TC_A(j,n,c) * R_A(j,k,n,c) * TF_A(j,n,c)$$
$$+ \alpha_3 \sum_{i}^{L} (TC_A(i,n,c) * R_A(i,k,n,c) * TF_A(i,n,c)$$

Composite trust metric in [2] (1)

Besides, this approach may also provide an update method to progressively modify the credibility of witnesses based on actual interaction performance (T_{actual}). Specifically, if their differences exceed a predetermined error threshold (ε), the reputation providers lose the credibility accordingly. Otherwise, they are rewarded by positively reinforcing their WTV. Clearly, in order to prevent cyclic fraud, they should be penalised in a greater amount than they are rewarded (Equations 2 and 3).

$$E_i = T_{actual}(k,n,c) - R_A(i,k,n,c)$$

The differences between actual interaction result and the witness's provided opinion (2)

$$TC'(i,n,c) = \eta * TC(i,n,c) + (1-\eta) * (adj) * 5$$
$$(adj) = +1 \; for \; |E_i| \leq \varepsilon$$
$$(adj) = -1 \; for \; |E_i| \geq \varepsilon$$

Adjusting the WTV (3)

Belief Model (R3b):Dempster-shafer theory of evidence (DST) is an extension to probability theory with the advantage of being able to model uncertainty. It is a widely used model which provides means for approximate reasoning under uncertainty. According to DST, there is no direct relationship between a hypothesis and its negation and as a result the summation of probabilities of atomic elements may not necessarily result into one. In this case, the remaining probability is interpreted as a state of uncertainty [9], [35].
The notion of DST is presented as follows:

***Definition 1*: Basic Probability Assignment:**
The basic probability assignment, denoted *bpa* or *m*, describes a mapping of the set of possibilities, denoted, *frame of discernment* Θ, to a value between [0, 1] where the function *m* is: $(1)\ m(\phi) = 0$,*and* $(2)\sum_{\hat{A}\subseteq\Theta} m(\hat{A}) = 1$

***Definition 2*: Belief Function:**
The belief function, denoted *bel* ($\hat{A}$) for a set $\hat{A}$ which $\hat{A} \subseteq \Theta$, is described as the sum of all the basic probability assignments over all proper subsets of $\hat{A}$. In the other word: $Bel(\{T, \neg T\}) = m(\{T\}) + m(\{\neg T\}) + m(\{T, \neg T\}) = 1$.
For example, if $m(\{T\}) = 0.7$,and $m(\{\neg T\}) = 0$ then $m(\{T, \neg T\}) = 0.3$

Suppose that, DST is employed as the underlying computational framework for T&R systems .In this case, the set of possibilities for each participant with regard to others can be defined as: $\Theta = \{T, \neg T\}$where *T* means that a given participant is trustworthy. To calculate trustworthiness of target agent A_j, agent $A_i$obtains information from two components: 1) direct experiences and 2) witness observations. With the help of DST, different methods could be proposed to calculate trust value for each component. For instance, [34]-[35] adopted upper bound (Ω_i) and lower bound (ω_i) thresholds from [18] to calculate trustworthiness value based on direct experience. Thus, the trustworthiness of A_j - $m(\{T\})$is evaluated as the summation of probabilities that satisfactory result $x_k$is obtained such that the domain of index variable (x_k) is initiated with upper bound threshold (Ω_i) and end with one.

On the other hand, in case A_i does not have an adequate amount of personal experience, [34]-[35] disseminate queries to obtain third-parties' testimonies regarding A_j. Afterwards, the received evidences are combined using *Dempster's rule of combination* [3], [28] to compute $m(\{T\})$based on the combined evidences. Within the context of

Journal of Theoretical and Applied Electronic Commerce Research
ISSN 0718–1876 Electronic Version
VOL 5 / ISSUE 2 / AUGUST 2010 / 97-117

This paper is available online at
www.jtaer.com
DOI: 10.4067/S0718-18762010000200007

T&R systems, since lack of evidences is easily expressed in belief model, it can be used as a means for differentiating between participants having uncertain reputation and bad reputation.

Fuzzy Model (R3c): Within fuzzy systems, the common parameters of T&R systems can be characterized via fuzzy sets. For instance, the quality of provided services, the accuracy of witness information, the reliability of opinions and time factor would be represented using linguistic terms such as *bad*, *average*, *good* for the first three factors and *short*, *average*, *long* for the time factor. Meanwhile, the membership functions describe in what degree we consider the QoS as *bad* or *good*. Furthermore, some trust models such as REGRET use fuzzy rules to reason about the reliability and trustworthiness of neighbors in providing accurate and honest opinions. For example, by demonstrating the social relationships among members of a community in the form of fuzzy rules, the degree of reliability of the participants can be simply determined.

Bayesian Approach (R3d): This approach uses the probability theory to model trust in T&R systems. In Bayesian systems, global reputation score is presented by beta Probability Density Function (PDF) with the shape parameters (α, β) which indicate the historical interaction results released by service consumers regarding a certain SP [9], [29]. In fact, within trust and reputation context, beta distribution would be interpreted as the posterior probability that a trustee agent fulfills its commitment towards the trusting agent in light of the previous outcomes experienced by witnesses or the trusting agent itself (Equation 4). Specifically, in some particular T&R systems, such as TRAVOS and BRS, that employ single rating systems, observations are recorded as frequency of *successful* and *failed* outcomes during certain time intervals; which are used to initialize the shape parameters α, β (Equation 5) [8], [19].

$$\rho(\tau|\alpha,\beta) = \frac{\gamma(\alpha+\beta)}{\gamma(\alpha)\gamma(\beta)}\,\tau^{\alpha-1}(1-\tau)^{\beta-1}$$

$$\text{where } \gamma(x+1) = x\gamma(x) \text{ and } 0 \le \rho \le 1, \alpha, \beta > 0$$

General formula for beta distribution (4)

$$\alpha = 0_s + 1 \quad , \quad \beta = 0_f + 1$$

Initialization of the shape parameters (5)

For instance, the following diagram displays the beta distribution with 26 successful and 4 failed outcomes (Figure 2). As is illustrated, the most likely value of trust happens in curve maximum. Thus, we can define the reputation score as a function of *expected value* as depicted in Equation (6).

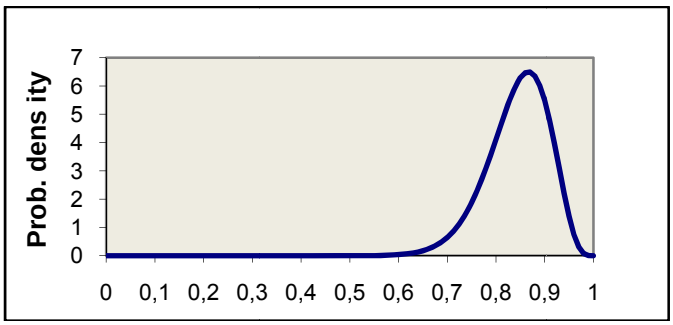

Figure 2: It Implies the Probability that SP Fulfills its Commitment is 84%

$$E(\tau) = \frac{\alpha}{\alpha + \beta}$$

The expected value of beta distribution (6)

3.1.4 Information Sources

Certainly, one of the promising characteristics of a T&R system is being able to provide reputation information in any condition. These systems enable participants to deduce trust value either from their ***direct experiences (R4a)*** or the ***witness's observations (R4b)*** information components. Furthermore, some T&R systems designed particular information components in response to the situation when neither of aforementioned information sources available. For example FIRE [6] has developed a ***certified reputation (R4c)*** component to perform actively for newly joined agents (both buyer and seller agents). In details, target agents release information indicating their former interaction results. The controversial problem that relates to this component is that; even if we assume that the target agent is honest and reliable in providing this reputation information there still is a probability that the target agent reveals only

This paper is available online at
www.jtaer.com
DOI: 10.4067/S0718-18762010000200007

the best results of its previous interactions and conceals unsatisfactory results. The other important information source which has been proved to be practical in MAS is ***role-based trust (R4d)*** [6], [20], [27]. As the name suggests, in this information source agents trust each other based on the predefined roles and relationships that exist amongst them. For instance, in case a seller agent belongs to the government or it has been certified by a trustworthy authority, buyer agents trust it with a high confidence level. Note that the role-based information source is updated as time progresses and may modify the relationships as a result of unexpected outcomes.

3.2 Soft Features

The soft features of T&R systems can be categorised along the following dimensions:

3.2.1 Context Diversity Checking

Aside from the reliability assessment of advisers in providing the reputation information, relying agents should be provided with ability to estimate the analogy of their negotiated contexts with its own potential ones intended to be experienced by the certain SP. On the other hand, the honesty and reliability assessment metrics would not sufficiently ensure the high quality service delivery unless it is accompanied by context diversity checking. In the following we have divided this dimension into two inter-related features:
Context similarity rate (R5a):As mentioned before, T&R systems are responsible for predicting the trustworthiness of potential SPs in providing agreed quality of service in a given context. Thus, any relying agent who issues the reputation query should specifically inquire about the trustworthiness of an SP in a particular context such as "*storing the fragile goods with Co1*". To illustrate, suppose that a very trustworthy recommendation agent (R-agent) had a direct experience with a queried SP in a rather different context ("*refrigerating goods with Co1*"). Therefore, despite of its high trustworthiness the exposed rating is not considered highly reliable. On the other hand, the opinion of a slightly less trustworthy R-agent who had direct interaction with the queried SP in a similar context should have more influence in the decision making process for the relying agent. To address this issue, T&R systems should develop a method to perform similarity checking between contexts in order to determine to what extent the received reputation values should be taken into account. The produced amount is employed in function which is designed to measure the influence degree of R-agents in trustworthiness computation.

Criteria similarity rate (R5b): Generally speaking, any communicating context between relying agents and specific SPs consists of several criterions such that the relying agent evaluates the performance and QoS based on them. Moreover, each criterion may have different influence and weight in the relying agent's perspective [2]. Hence, the measured trustworthiness value is mainly dependant on how much these criteria are fulfilled.

Lets suppose that two matched contexts with the same set of criteria come with different weights assigned by corresponding agents *A* and *B*. As it depicted in Table (2) and Table (3) even though the contexts are perfectly matched and certain SP delivered same quality of service, the relying agents experience rather different QoS owing to the different influence value ascribed to each criterion by the relying agents. Alternatively, if these two agents act as R-agents, their recommendations would not be beneficial unless they are aware of the criterion's influences in the relying agent's viewpoint. To clarify, if the R-agents elicit the preferences of the criteria, they might predict that unlike their cases, the relying agent will be fully satisfied with the delivered service (Table 4). Note that the context diversity checking should be executed right after the reliability and honesty assessment of R-agents when we are rather confident regarding the accuracy and honesty of the received ratings.

3.2.2 Adaptability

Efficient T&R systems should take an adaptive approach to deal with the inherent dynamism characterizing the open environments where they operate. For example, they should provide appropriate methods to perform effectively in a scalable environment and address the problem of newcomers who intend to establish mutual trust relationships with others. In this subsection some of the common issues of T&R operation in open dynamic environment are being discussed.

Table 2: Relying agent *A* assigns a trustworthiness of 53% to service provider *S1* in the context of "online shopping service"

Criterion	Importance weight	Actual delivered service	Expected delivered Service
Responsive customer service	2	10	10
Intact delivery	10	10	10
On time delivery	6	6	10
Agreed cost	10	1	10
Money back guarantee	5	1	10
		171	320
		53%	

Journal of Theoretical and Applied Electronic Commerce Research
ISSN 0718–1876 Electronic Version
VOL 5 / ISSUE 2 / AUGUST 2010 / 97-117

This paper is available online at
www.jtaer.com
DOI: 10.4067/S0718-18762010000200007

Table 3: Relying agent *B* assigns trustworthiness of 43% to service provider *S1* in the context of "online shopping service"

Criterion	Importance weight	Actual delivered service	Expected delivered Service
Responsive customer service	2	10	10
Intact delivery	10	10	10
On time delivery	3	6	10
Agreed cost	8	1	10
Money back guarantee	10	1	10
		153	350
		43%	

Table 4: Predicted trustworthiness of 76% calculated by agents *A*, *B* after eliciting the preferences of the relying agent

Criterion	Importance weight	Predicted delivered service	Expected delivered Service
Responsive customer service	10	10	10
Intact delivery	10	10	10
On time delivery	7	6	10
Agreed cost	2	1	10
Money back guarantee	4	1	10
		254	330
		76%	

Bootstrapping (R6a): Due to the openness of MAS any participants (sellers and buyers) could dynamically join and leave the system. As time passes, the participants get progressively acquainted with the environment thereby establishing a connection with reliable and trustworthy neighbors and enhancing their reputation values. However the issue arises when newcomers with no acquaintance join the community with fairly low initial reputation values [5]. For instance, In case of newly joined SP, source agents are usually reluctant to initiate communication with this SP even though it may offer reasonable or even better service compared to the other available ones. On the other hand, when the newcomers are source agents who are willing to exploit the particular service, SPs might decide not to take a risk and establish connection with them. Note that despite the advantageous feature of bootstrapping in hindering malicious agents from changing their identity and re-entering the system [16], it may entail the exclusion of newcomers from engaging in the open MAS. Thus, T&R systems should think about strategies to address the bootstrapping issue and not simply ignore such problems.

Dynamism in open MAS (R6b):In a dynamic environment, it is impossible to predict all the forthcoming incidents in advance. Thus, any T&R system should be equipped with techniques to deal with unanticipated events such as changes in participant populations and attitudes. Moreover, since in open MAS it is quite probable that some information sources would not be temporarily available; conditions should be created for any participants to be able to evaluate the performance of candidate SPs ubiquitously at any time. In addition, in order to operate effectively under any circumstances, T&R systems might provide mechanisms to monitor the behavior and relationships of all participants, including the SPs and witnesses, and thus learn and update respective information correspondingly.

3.2.3 Reliability & Honesty Measurement

In this dimension we have described some particular features which are required to increase the reliability and precision of the generated trustworthiness value. In fact, with the employment of these features we are able to alleviate the effect of dishonest information providers and spurious ratings. We have classified this dimension into four main aspects:

Trend and volatility detection in SP behavior (R7a): In reputation prediction and measurement, it is beneficial not only know the reputation value of a SP at certain time interval but also its trend line and behavior model in a last few time intervals [2], [27]. Moreover, by observing the trend line of SP, we could simply detect its volatile and periodical behavior as a result of fraudulent activity or its incapability. To clarify, it could occur that a particular SP provides satisfactory quality of service in most situations when there is not much at stake whereas it acts conversely in some occasions associated with a large gain. In this case, if we overlook the trend factor in reputation prediction and simply average the reputation values, the occasional fraud is likely to be masked (Figure 3). Furthermore, if we observe variations in the quality of service in the same time interval, we can deduce that a certain service provider is incapable to handle high workload being experienced and not easily judge it as deceptive and untrustworthy.

Journal of Theoretical and Applied Electronic Commerce Research
ISSN 0718–1876 Electronic Version
VOL 5 / ISSUE 2 / AUGUST 2010 / 97-117

This paper is available online at
www.jtaer.com
DOI: 10.4067/S0718-18762010000200007

Besides, by knowing the overall trend, we can evidently figure out if a particular SP is deteriorating over time (Figure 4). One possible solution to deal with this problem is by providing reliability measurement which helps to calculate the confidence level of the generated expected reputation value. However, even though the inconsistency and QoS variation of SP is being discovered with the help of reliability measures the occasional fraud and periodical behavior still remains undetected.

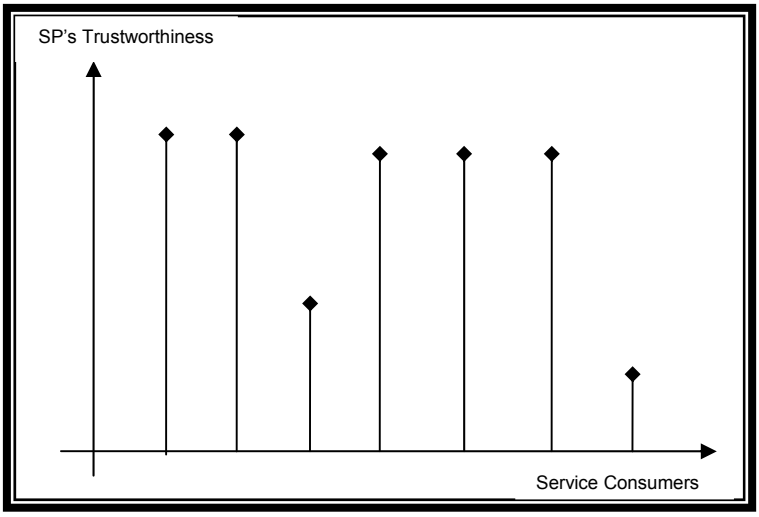

Figure 3: Detecting fraudulent activity in SP's behavior model confronting various service consumers

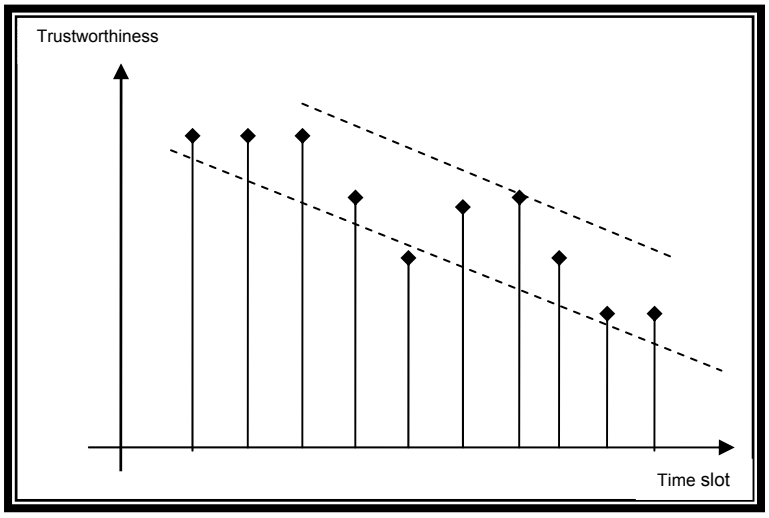

Figure 4: Deteriorating trustworthiness value with time [2]

Proficiency of R-agent in giving correct ratings (R7b): the other important parameter which should not be neglected in reliability and honesty measurement is related to the competency of an R-agent in giving correct ratings. In this parameter we are concerned about the proficiency and *eligibility* of an R-agent in publishing correct ratings rather than the *reliability* of an R-agent in giving honest ratings. For example, the ratings might be honest however they might suffer poor judgment or insufficient interactions with the SP. Even though in this case the R-agent is reliable, it lacks the competence in publishing true ratings.

Ballot box stuffing (R7c) (aka double counting or the correlated evidence problem): the characteristic of open MAS enables varied kinds of agents including malicious, anonymous, trustworthy and untrustworthy agents to actively participate and perform in the environment. Therefore, in order to address the lack of security and trust in open MAS, some known agents may form a chain of trust relationship to provide a transitive trust among members. This coalition can act positively by sharing experiences and ratings as well as providing group opinions upon request; however, it excludes other agents who they may not like or know. This matter could lead to ballot box stuffing where the community of agents attack competitors by issuing unfair ratings and recommendations against them. To explain, the coalition would cast negative votes against the outsiders and alternatively cast positive votes in favour of its members [9]. Thus, T&R systems should employ mechanisms to encounter the ballot box stuffing problem while supporting the advantageous features of agent's coalitions. One of the potential solutions to address this problem could be defining specific thresholds which dynamically determine the maximum number of eligible votes as well as

Journal of Theoretical and Applied Electronic Commerce Research
ISSN 0718–1876 Electronic Version
VOL 5 / ISSUE 2 / AUGUST 2010 / 97-117

This paper is available online at
www.jtaer.com
DOI: 10.4067/S0718-18762010000200007

providing the controlling mechanism which restricts participants to release *only* their self experiences and not the other's opinion.

Reliability of R-agent in giving honest ratings (R7d): it is quite possible that R-agents provide misleading ratings regarding queried SP owing to jealousy, competitive or monetary reasons even though in their own dealing with the relying agent as a SP they might have constantly satisfactorily fulfilled the commitment in any agreement. In this case, R-agent is considered to be credible and trustworthy as service provider but not reliable in the role of recommendation agent. Thus, in order to accurately predict and measure the reputation of particular SP, any T&R systems should utilize mechanisms to determine the credibility level of ratings by detecting intentional errors and unfair ratings rather than simply labeling any known R-agent as reliable.

Discrimination detection (R7e): In open MAS, it is probable that specific SP perfectly meets a service level agreement for the majority of relying agents but excluding a small subset of them. Likewise, it might happen that a particular SP offers preferential service to a selected relying agent while treating the others differently. These types of discriminations lead to unfairly judging the honesty of advisers who are distinguishably treated by the queried SP, since their reputation values are conflicting with the rest. Moreover, an adviser can also play the role of discriminator such that it provides fair ratings except when dealing with a specific SP [9], [27].Thus, T&R systems should exploit techniques to detect the above mentioned discriminations and differentiate between victim and dishonest R-agents.

3.2.4 Basic reputation measurement parameters

This dimension addresses some crucial parameters which may increase the accuracy of the expected reputation value.

Transitivity Rate (R8a): To accurately measure the expected trustworthiness of a particular service provider (SP), it is required to rank the incoming recommendations issued by various knowledgeable parties. More explicitly, any T&R systems should enable its participating agent to assign more value and weight for the recommendations coming from known parties who s/he has already had interaction with and consider its recommendation as first-hand rating (with the transitivity degree of one) which have a significant influence in the decision making of the relying agent. Similarly, when the relying agent issues a reputation query regarding a particular SP in MAS, it may happen that numerous participants respond whose trustworthiness levels are unknown to the relying agent. Their recommendations are considered as third-hand ratings (transitivity degree of three) and may have the least influence on the trustworthiness measurement. Moreover, in MAS, it is quite probable that known parties have not had any records of interactions with the queried SP; however, they respond with a list of agents who had experience with a queried SP. Clearly, their recommendations have an intermediate effect in the determination of expected reputation value (depending on the level of trustworthiness of their recommended party) and are considered as second-hand ratings (with transitivity degree of two). It is noteworthy to mention that if the relying agent had direct interactions with a certain SP, this information is thought to be the most important and the relative calculated trustworthiness value is considered to have significant impact in the decision making phase .It is also capable to *veto* the trustworthiness value obtained from the other information sources.

Time (R8b): time should be considered as one of the essential parameters in any reputation measurement which indicates the recency and freshness of the ratings revealed either by R-agents or the ones recorded in the relying agent's database. In this case, the older the ratings are, the less influence they have on the reputation and trustworthiness calculation. One of the reason that signifies the importance of the time factor stems from the changes which may occur in ownership and management of SPs as time progresses, leading to different internal policies in service provisioning [2]. For this reason, ratings before a certain time interval should not be evaluated.

4 Comparing T&R Systems using the proposed Framework

In this section, we intend to examine the effectiveness of the trust and reputation systems described in Section 2 by providing an argumentative comparison across the framework. Table (5) provides a comparison of the reviewed T&R systems against the refined features presented in Figure (1).

The meanings of symbols used in Table (5) are as follows:

- **N/S** the model does not satisfy the corresponding feature.
- **P**: the model attempts to address corresponding feature and has partly succeed.
- **Y**: the model satisfied the corresponding feature.
- **A**: the model assumes the particular feature exists and does not provide any method to address it.
- **N/A**: the corresponding requirement is not applicable.

This paper is available online at
www.jtaer.com
DOI: 10.4067/S0718-18762010000200007

4.1 FIRE Model

The decentralized FIRE framework is developed to deal with the dynamic characteristics of open MAS (R6a). To name a few, FIRE is able to handle potential problems in open MAS such as scalability (due to openness of system) and changeability in participant's behaviors or relationships (for example, when a former trustworthy partner becomes unreliable or agents may break the old relationships and make new ones depending on their goals or situation). In order to maintain an effective operation under such circumstances, it continuously monitors the performance of components and adopts learning techniques with the purpose of adjusting respective parameters tailored to the current situation [5]. It is also a generic model which can be instantiated and applied in a wide range of applications. Moreover, with the help of various information sources, FIRE is able to effectively deduce trust in almost any situation; it is even able to address the bootstrapping problem when a particular service provider has newly joined the system (R6b). Furthermore, in order to deal with the malicious third-parties who provide misleading reports, FIRE employs a credibility model to assess the honesty of the revealing reports and consequently filters out the lying reporters (R7d). Yet, suppose that a particular agent undergoes preferential treatment or discriminational behavior cases in which its ratings do not reflect the actual performance of the SP. Despite its creative attempt to differentiate between a dishonest and inaccurate report from an honest but wrong one, FIRE cannot fully satisfy this requirement (R7b, R7e). To clarify, in the credibility model it defines a threshold that indicates the maximal acceptable differences of the exposed reputation values with the actual interaction result. Therefore, any reporter whose inaccuracy exceeds the threshold is labelled as dishonest and is heavily penalised by losing its credibility value; clearly, an honest but mistaken agent is no exception. The possible solution to this problem would be to tune the threshold to a higher value to reduce the probability of falsely classifying the honest witnesses [5]. However, this results in delaying the process of discarding dishonest witnesses. To resolve this concern, FIRE exploits techniques to automatically tune thresholds according to the performance deviation of SPs (based on direct observations of evaluator agents); but, still there will be a trade-off in distinguishing deceptive third-parties versus mistaken ones.

The other significant feature of this framework is its ability to measure the reliability of an SP in providing the expected level of trustworthiness. Although, its deviation and rating reliability are well-suited to detect frequent fluctuations in the service provider behavior, FIRE is unable to detect the deceitful activity of the SP in cases when it can obtain a large profit (R7a). Moreover, FIRE cannot perceive the overall behavioral trend of an SP over time. To explain, if an evaluator agent becomes aware of the deteriorating trend of an SP as time passes, it underrates the expected trust value and feels pessimistic towards the future interaction results.

As mentioned before, FIRE employs a multi-criteria rating system such that the participants rate the performance of SPs based on predefined criteria rather than providing an overall opinion; however, it does not provide techniques to perform context diversity checking whereupon the context and criteria similarity rates are not calculated (R5a, R5b). Thereby, it is unable to elicit the preferences of relying agents in order to predict the trust value of a particular SP more precisely.

4.2 REGRET

The REGRET T&R system takes advantage of a variety of information components to predict the trustworthiness of target SPs almost in any situation. Distinctively, in order to make more accurate judgements, it provides the *neighbourhood and system* reputation components in addition to the direct interaction and witness reputation components.

Using social relationships, it enables newcomers to take part in the community's activities; thus provides the possibility for them to increase their knowledge and improve their social status persistently (R6a) [27]. Moreover, due to the dynamic characteristic of an open environment, the population of participants varies from time to time. Besides, the agent's behavior and performances oscillate, being influenced by unexpected changes in such environments. Evidently, REGRET is incapable to extensively deal with the dynamicity of an open MAS thus cannot perform effectively under all circumstances of such an environment (R6b) [5].

As aforementioned, the distinguishing feature of REGRET is its use of social relationships between participants in modeling trust. With the help of the defined social relations, source agents are able to identify suitable witnesses and provide appropriate recommendations with regards to a target agent. Furthermore, REGRET proposes a mechanism to handle the ballot box stuffing and correlated evidence problem where set of witnesses express their opinions based on the same experiences. To do this, it groups the potential witnesses and considers each of them as individual sources of information and then uses a heuristics to select the best representative in the group to send the query to. This matter reduces the number of sent queries as well as alleviates the effect of correlated evidence problem (R7c). However, REGRET assumes that each agent owns pre-defined sociograms which display social relationships [26] and does not address how to locate witnesses in these social structures (R2).

Subsequently, in order to ascertain the quality of the provided recommendations, any T&R systems should develop techniques to detect deceptive and unreliable agents and following that underrate their reputation values or ignore them, accordingly. For this purpose, REGRET mainly relies on social relations and states them via fuzzy rules. Through these rules, it validates the obtained recommendations and determines their influence degree in the

This paper is available online at
www.jtaer.com
DOI: 10.4067/S0718-18762010000200007

reputation aggregation method. It is noteworthy to mention that REGRET examines the truthfulness of information in general and does not differentiate between dishonest third-parties and incompetent but honest ones (R7b), (R7d). Moreover, it does not describe any method to update the weight parameter of witnesses after actual interaction results have been obtained.

The other significant characteristic of this T&R system is its ability to support multiple-criterion rating system (R2a). As already mentioned in Section 2, REGRET is aspect-oriented and records reputations linked to a single behavioural aspects of a contract [24]. For instance, in case the contract consists of multiple criteria, it specifically inquires regarding particular criterion rather than a general reputation value. Alternatively, in order to calculate overall ratings, REGRET enables each participant to design an ontological structure of the contract suited to its requirement and weights each aspect proportionally. This feature addresses the criteria similarity rate (R5a) in a manner which leads to more precise prediction of the reputation value.

4.3 T&R Model by Yu and Singh

Yu and Singh have proposed a decentralized reputation management model to locate the rightful witnesses in MAS in order to evaluate the trustworthiness of an SP which is willing to communicate (R2a). However, it's recommended Trust Network is not capable of distinguishing between honest and deceptive agents. To deal with this problem, a deception model was developed to detect spurious ratings as well as updating the behavioral models of acquaintances based on their performance (R7d). Moreover, to avoid double counting and correlated evidence problems, it restricts participants to release only their local belief; however, it does not analyze the possible solution for the decentralized environment without central authority (R7c).

Furthermore, even though it defines two thresholds as upper and lower bounds of trust and also counts on the state of uncertainty in calculating the expected trustworthiness value, this model does not provide any reliability measures to figure out the confidence level of the produced value (R7a).

Finally, Yu and Singh have attempted to design a reputation model compatible with the inherent dynamicity in open MAS. For example in their approach, individuals can dynamically choose their neighbours from their current acquaintances. In addition, when the majority of agents exhibit volatile and changing behavior, it can adaptively adjust the exaggerative coefficient to a higher value to swiftly filter out deceitful agents from the system (R6b).

4.4 TRAVOS

TRAVOS is a probabilistic trust model which uses beta distribution probability functions to calculate the likelihood of certain SPs fulfilling agreed obligations given its past personal experiences and reputation information (R3d).

Even though the performance of TRAVOS is validated in a decentralized online marketplace with pre-determined agent populations, it can extend easily to large scale open systems [29]. Yet, it lacks the ability to address the bootstrapping problem as well as the dynamicity in participants' behavior which may change their attitude overtime [5] (R6a, R6b). Besides, this model assumes that reputation information is accessible upon demand and does not present any approach for locating witnesses (R2).

Using probability theory, TRAOVS provides a novel approach for detecting and filtering malevolent witnesses. It adjusts the effect of provided opinions on the trustworthiness measurement; corresponding to the accuracy degree of their reporters. However, it does not provide any reliability measure to assess the degree of confidence of a truster agent in achieving the expected performance from the trustee.

Moreover, this model attempts to tackle discrimination and ballot box stuffing problems by underrating exaggerative opinions (R7c), (R7e); however, it is unable to detect fraudulent participants who lie in small amounts. Furthermore, since TRAOVS is based on a single rating system such that the reputation is shared in the form of frequency of successful and unsuccessful interaction results, it is incapable of providing suitable recommendations in an environment with competitive service providers offering variety of services in different contexts (R5).

It is noteworthy to mention that, this T&R system is mostly comparable with BRS in the context of handling inaccurate reports [8], [31]. Nevertheless, BRS is based on an endogenous approach which presumes that the majority of reputation sources provide an accurate opinion thus discards any opinions that deviate considerably from the average [19]. Moreover, unlike TRAVOS which significantly relies on personal experiences (section 2.4) BRS does not differentiate between direct experiences and reputation information and treats them equally in trustworthiness computation.

4.5 PeerTrust

The PeerTrust T&R system consists of various distinguishing features which helps it to significantly outperform a number of available T&R systems in some particular contexts. Specifically, it puts special effort to tackle problems relating to the reliability and honesty assessment (R7). For instance, by means of the credibility measures it is able to

This paper is available online at
www.jtaer.com
DOI: 10.4067/S0718-18762010000200007

act effectively against malicious coalitions thereby addressing ballot box stuffing and correlated evidence problem (R7c) [33]. Furthermore, the use of a transaction context factor helps in revealing actual intention of opportunistic peers who cheat whenever it is advantageous for them to do so. In light of these transaction context parameters such as size and time as well as adaptive *time window-based* algorithm PeerTrust can detect volatile personality, oscillating attitude and discriminative behavior of participants and thus adjust their trust value accordingly (R7a), (R6b). (The basic idea is to adaptively calculate the trust value using smaller time window which reflect most recent behavior of peers and compare it with the pre-determined time window. If they differ more than a pre-defined threshold, this implicitly signifies the volatility and fraudulent attitude of peers.) However, it does not provide any reliability measures to estimate the confidence degree of the generated trust value.

The characteristics of peer-to-peer online communities oblige T&R systems to support dynamicity in order to act responsively in all unpredictable situations. For this purpose, PeerTrust provides means to calculate trustworthiness of target peers in *almost* all circumstances. For instance, when requesting peers do not have adequate information to evaluate the credibility of witnesses, this model proposes an alternative approach for credibility measurement. According to this approach, the requesting peer recursively considers other peers' trust values as a credibility factor in order to compute the target peer's performance [32]. Even though this technique provides initial judgement for target peer's trust value, still requesting peers are threatened with the risk of confronting misleading participants who disseminate distorted feedback. Furthermore, this decentralized scalable trust model does not address the bootstrapping problem of newly joined peers (R6a).

Notably, since this T&R system adopts a binary system which presents the interaction results as satisfactorily/unsatisfactorily, it is restricted to execute context diversity checking in order to provide the recommendation more analogous to requesting peer's expectations (R5).

4.6 BRS

BRS presents a set of rich features which differentiate it from some existing T&R systems in certain ways. In particular, it proposes a novel approach to rectify the bootstrapping problem of the newly joined agent (R6a). That is, this reputation system dynamically assigns a base rate reputation score to newcomers upon arrival. It provides a method to track the average reputation scores of the whole community so as to settle the newcomers into a conservative state. Notably, such base rate could have be biased towards either positive or negative reputation scores depending on the overall participants' trustworthiness attitudes and the quality of the market at the time. [11], [12].

Furthermore, BRS takes a step towards tackling the inherent dynamicity of an open marketplace (R6b). Unlike other available T&R models which mainly concentrate on modeling the *adviser's* behaviour, BRS models the behavioral pattern of buyer and seller as well. In particular, BRS provides sellers with the ability to adaptively change their behavior to increase their benefits while maintaining a satisfactory level of honesty. For instance, based on a set of heuristics, if a certain seller agent does not succeed in conducting any business for certain period of time, it will automatically decrease the selling price while increasing its level of honesty. On the other hand, it defines the risk attitude parameter for buyer agents which affects the purchasing pattern of the buyers. That is, if the buyer makes a large loss in previous interactions, it intelligently increases the risk-aversion parameter for next rounds of transactions [31].

Furthermore BRS provides a robust protection mechanism against both positive and negative unfair ratings. As such, to diminish the risk of malicious advisers who attempt to manipulate the reputation system for their own benefits, it provides statistical iterated filtering techniques based on beta distribution to dynamically expel such advisers with unsatisfactory rating levels. In so doing, it defines an adaptive *quantile* parameter q, which is the point at which 1% of the ratings fall below and 99% of ratings fall above that value [31]. With the appropriate adjustment of this parameter and the *longevity* factorλ, the filtering algorithm can perform more efficiently in excluding of old and unfair ratings in certain conditions. That is, empirical results indicate that [31], the basis of the filtering algorithm enables BRS to efficiently detect the ballet box stuffing problem (R7c) as well as restraining deceptive advisers to distort seller's reputation when the majority of participants provide honest ratings. In contrast, when a substantial proportion of participants is dishonest (over 30%), this technique may mistakenly reject the correct ratings which results in poor estimation of the trustworthiness of the target seller. Moreover, BRS can successfully detect the volatility in the agents' behavior in case malevolent agents choose a strategy to intersperse unfair ratings with fair ratings over 50% of the time (R7a).

Finally, by the means of supporting the continuous ratings input, it not only enhances applicability and flexibility of BRS in dealing with the continuous nature of some observations; it also increases the reliability and confidence degree of the provided ratings.

Journal of Theoretical and Applied Electronic Commerce Research
ISSN 0718–1876 Electronic Version
VOL 5 / ISSUE 2 / AUGUST 2010 / 97-117

This paper is available online at
www.jtaer.com
DOI: 10.4067/S0718-18762010000200007

Table 5: Comparing the reviewed T&R systems across the dimensions of the framework.

	R1		R2			R3				R4				R5		R6		R7					R8	
	a	***b***	***a***	***b***	***c***	***a***	***b***	***c***	***d***	***a***	***b***	***c***	***d***	***a***	***b***	***a***	***b***	***a***	***b***	***c***	***d***	***e***	***a***	***B***
FIRE		Y	Y			Y				Y	Y	Y	Y			Y	Y	P	P	Y	Y	P	Y	P^a
REGRET		Y	A					Y		Y	Y		Y	Y		Y				Y	Y		Y	
Model by Yu & Singh			Y				Y			Y	Y					P^b	Y			P	Y		Y	
TRAVOS	Y		A						Y	Y	Y			N/A	N/A					Y	Y	P^c	Y	
PeerTrust	Y		Y			Y				Y	Y			N/A	N/A	Y	Y	P^d		Y	Y	P^c	Y	
BRS	Y^e		A^f						Y	Y	Y			N/A	N/A	Y	Y	P^g		P^h		P^c	Y	

P^a: FIRE [6] model , implicitly incorporates *transitivity rate* by means of providing suitable trust metric along with reliability measures for each individual information sources.

P^b: This model treats agents with bad reputations and newcomers -with no reputation- differently.

P^c: This T&R system may detect discrimination behavior but cannot differentiate between dishonest and victim R-agents.

P^d: There is no reliability measure.

Y^e: This reputation system supports multinomial and binomial models of ratings. It also supports both discrete and continuous ratings as input.

A^f: This reputation system is compatible with all three witness location approaches.

P^g : this system is able to detect the trend and volatility in participants behavior if they play the unfairness strategy with probability above 0.5.

P^h: BRS cannot effectively combat the ballet box stuffing problem when the majority of participants provide unfair ratings.
The presented T&R systems encompass important features in order to bring robustness within a virtual community. They exploit ad-hoc techniques to combat threats and vulnerabilities inherent in the particular environment where they have been deployed.

Table(5) highlights the differences between the selected T&R systems and compares them across the dimensions of the framework. As can be noticed, some of the presented T&R systems addresses a wider range of dimensions whereas the others cover a smaller number of them. This matter would not necessarily imply better-quality and applicability of such systems compared with the others. Instead, one should consider the context in which these T&R systems are employed and evaluate how well they accomplish the goals and requirements of that particular environment.

In other words, the nature of some communities necessitates the implementation of some precautions and defense mechanisms to curtail attacks that exploit the intrinsic vulnerabilities of such communities. For example, the T&R system designed for an open anonymous auction marketplace should be resistant to re-entry attack where as for trust model designed for the environment with identity management service this is not a challenge. That is, the T&R systems which are robust in one community could be vulnerable in another community [10].

As aforementioned, each of the presented systems exhibits a particular quality well-suited to their application domains and the related requirements. For example, FIRE is relatively robust in open scalable multi-agent environments like electronic marketplaces where a variety of anonymous agents with changing behaviors participate.

This paper is available online at
www.jtaer.com
DOI: 10.4067/S0718-18762010000200007

Using the advanced reliability assessment techniques, this model is able to deal with malicious agents and expels them from the system in a timely manner. However, its reliability measure seems to be vulnerable in case *majority* of advisers provide unfair ratings. With exploitation of different information sources, the bootstrapping problem is well-addressed in FIRE to facilitate the establishment of mutual relationships between participants. This matter makes FIRE an appropriate model to be employed in *recommender systems* that face the bootstrapping problem. Moreover, since this model enables participants to carry out bilateral and multi-attributive negotiation over given merchandise; it can be appropriate for use in the market where service providers offer different qualities of services.

On the other hand, REGRET is well-established for*competitiveB2B e-commerce* environments where each of the business partners could be equipped with an individual sociogram to locate the most trustworthy and qualified transaction partners. Through the identification of social relationships, REGRET is able to model the cooperativeness and competitiveness attitudes of participants which could be utilized for evaluation of their feedback's credibility. However, this model is well-suited for such market types consisting of participants with consistent behavioral patterns as it indicates vulnerabilities where the market participants oscillate their behavior dynamically. In addition, the use of REGRET in B2B applications facilitates negotiation of items with arbitrary variable attributes which increase the variation of products offered by service providers.

The scalable and dynamic Trust Network makes Yu&Singh T&R model suitable for the *open decentralized* environments where the number of participants is changing constantly. Its credibility measurement approach provides the means to model the behavioral patterns of agents and deal with the strategic dynamic personality of participants. The Yu&Singh T&R model considers a wide range of behavioral attitudes of participants including dishonest with complementary pattern, fairly honest with exaggerative negative or positive inclination; and fully honest pattern. In fact, such model indicates its usefulness in revealing the actual intention of participants and adjusting their influence degree accordingly. This feature of Yu&Singh T&R model underlines its robustness in the environments where the *majority* of dishonest participants with diverse behavioral patterns take over the community.

The scarcity of trust becomes a huge bottleneck for online auction markets. In particular, this environment is vulnerable against a set of attacks such as collusion and ballot box stuffing which can significantly manipulate any trust and reputation systems. The TRAVOS T&R system builds up certain features to promote trust in *online auctions*. By providing an adaptive filtering technique that evaluates honesty of the participants with respect to their previous performance, it is able to mitigate the bidder collusion affect and their exaggerative opinions. Furthermore, one of the daunting challenges in online auctions is the condition when a lying auctioneer places a shill in the market to provide discriminative bidding. To rectify this, TRAVOS provides the method to under rate such biddings so as to make them well-matched with the rest of the biddings. With regards to its single rating approach, even though it seems sufficient to bid over the general quality of products deploying multi-criterion rating approach brings more flexibility for participants to make better decisions considering several aspects.

The architecture of PeerTrust makes it compatible with several sizes of the*peer2peer communities*. In P2P file sharing systems, the criteria of the agreement are implicitly known by peers, thus the employed T&R systems like PeerTrust does not require the means to perform context diversity checking. Furthermore, P2P systems are vulnerable against various attacks such as collusions and ballot stuffing which adversely affect the robustness of the community. Thereby, through the notion of credibility factor, this model developed a defensive mechanism to act strongly against malicious coalitions aiming at manipulating the reputation of community's members. In addition, PeerTrust intuitively stimulates the cooperativeness attitude of its participants. This feature seems necessary for such environment where members indicate least motivation to collaborate with others in providing the reputation information. In light of the community context factor, PeerTrust provides incentive for self-interested agents to disseminate their feedback regarding the target peer.

The single ratings approach makes BRS very suitable in *online blog and news communities* where members can post articles and anyone can subjectively rate them. Specifically, by providing the multinomial rating style, BRS enables its users of the community to express their genuine opinions more delicately across a set of predefined ratings levels. Furthermore, BRS exploits the adaptive filtering algorithm which is quite sensitive to the environmental circumstances and quality of participants. It can detect the volatility in a participant's behavior in case some deceptive members strategically rate down other's comments to promote their own reputation scores. In addition, its endogenous filtering techniques provide robustness for such communities where a small portion of the participants provides discriminatory ratings. In spite of its excellent suitability in various aspects, however; BRS might display some vulnerability in such risky environments where the *majority* of misbehaved and noisy participants take over the community [31].

5 Concluding Remarks and Open Problems

In this paper, we have introduced a framework for classifying and comparing Trust and Reputation systems and provided an overview of some prominent current T&R systems according to this framework pointing to ways to choose one over another for particular applications. The dimensions of this framework help system-developers to choose or build their desired T&R system with appropriate features according to their requirements. The trust and reputation framework presented in this study covers a broad range of applications, as such not all dimensions might

This paper is available online at
www.jtaer.com
DOI: 10.4067/S0718-18762010000200007

be applicable in some particular application contexts. Therefore, to build a suitable T&R system one shall first identify the application constraints, specifications and what is desired to have in the system. Then, the appropriate features from the framework need to be selected to fulfill pre-determined expectations.

For instance, designing T&R systems for open online community and peer-to-peer systems requires a decentralized adaptive architecture with a context-dependent rating system that examines the recommended service providers more *personally* such that we will be confident that the selected service provider delivers its commitment as we anticipate. Furthermore, due to the anonymity of members in online communities, they may hide their actual intentions and opportunistically cheat whenever it is advantageous for them. Hence, in such environment picking the optimum and efficient reliability assessment techniques seems critical. Moreover, addressing the bootstrapping problem of newly-joined members should be one of the priorities of such open systems. However, they should choose mechanisms to distinguish between newcomers and malicious members who may disguise their identity and return to the system for more sabotaging activities. Besides this since dynamicity is an intrinsic feature of open environments, the T&R system should employ certain computation engines with qualification in predicting the trust value of peers in the absence of some information sources. Moreover, including basic parameters such as time, transitivity rate and context factors brings more advantages.

On the other side, in developing centralized T&R systems for online auction applications like eBay (Site 1), the availability of information sources is not an issue. More explicitly, since the reputation computation engine is taken over by the central authority, it can ubiquitously access the pool of information upon demand. Furthermore, the ballot box stuffing and correlated evidence problem is out of question in a centralized system. However, the main concern of such systems is with the lack of an effective rating system which could articulately assess the delivered services analogous with real-life experienced judgement. To support this perspective, [21] claims that provided feedback in eBay is unrealistically positive such that service providers receive negative rating only 1% of the time which indicates the incapability of such systems in providing informative ratings.

Understandably, there is no single solution appropriates for all kinds of applications and environments. In this paper, we attempt to provide the means to find the most appropriate path to examine the applicability and usefulness of the current T&R systems across different application domains. Following this, we have summarized the most common features of trust and reputation systems and described how the existing systems support these features.

Although there has been a significant number of works in T&R systems, there are still some open fields that need further explorations. Specifically, several work has been done on *reliability and honesty assessment* which proposed innovative solutions in dealing with spurious feedback in uncertain environments. However, some critical aspects of this feature are not fully supported in current trust and reputation systems. To name a few, addressing *discrimination detection*, *Trend & volatility detection in service provider behaviors*, *ballot box stuffing* and *distinguishing between malicious and victim participants* is not yet addressed thoroughly.

In addition, it remains a challenge to build an informative rating system which supports the *context diversity checking* feature by providing *context* and *criteria similarity rate* to considerably improve the quality of judgements and recommendations.

Websites List

Site 1: Online auction website
www.ebay.com

Site 2: Multinational Electronic Commerce shop
www.auctions.amazon.com

Site 3: Decentralized, censorship-resistant distributed data store
www.zeropaid.com\freenet

References

[1] D. Artz and Y. Gil, A survey of trust in computer science and the semantic web, Journal of Web Semantics, vol. 5, no. 2, pp. 58-71, 2007.
[2] E. Chang, T. S. Dillon, and F. Hussain, Trust and reputation for Service Oriented Environments-Technologies for Building Business Intelligence and Consumer Confidence. Location: US: John Wiley & Sons, 2006.
[3] J. Gordon and E. H. Shortliffe, The Dempster–Shafer theory of evidence, in Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project, in Buchanan and E. H. Shortliffe, eds. Addison Wesley,Reading,MA,1984, pp. 272-292.
[4] F. K. Hussain, E. Chang, and T. S. Dillonn, Trustworthiness and CCCI metrics in P2P communication, International Journal of Computer Systems Science and Engineering, vol. 19, no. 3, 2004.
[5] T. D. Huynh, Trust and Reputation in Open Multi-Agent Systems, PhD thesis, University of Southampton, 2006.

This paper is available online at
www.jtaer.com
DOI: 10.4067/S0718-18762010000200007

[6] T. D. Huynh, N. R. Jennings, and N. R. Shadbolt, An integrated trust and reputation model for open multi-agent systems, Journal of Autonomous Agents and Multi-Agent Systems, vol. 13, no. 2, pp. 119-154, 2006.
[7] T. D. Huynh, N. R. Jennings, and N. Shadbolt, Developing an integrated trust and reputation model for open multi-agent systems, in Proceedings of the 7th International Workshop on Trust in Agent Societies, New York, 2004, pp. 62–77.
[8] A. Jøsang, and R. Ismail, The beta reputation system, in Proceedings of the 15th Bled Conference on electronic Commerce, Bled, Slovenia. 2002.
[9] A. Jøsang, R. Ismail, and C. Boyd, A survey of trust and reputation systems for online service provision, Decision Support Systems, vol. 43, no. 2, pp. 618-644, 2007.
[10] A. Jøsang and J. Golbeck, Challenges for robust of trust and reputation systems, in Proceedings of the 5th International Workshop on Security and Trust Management(SMT 2009), Saint Malo, France, September 2009.
[11] A. Jøsang and W. Quattrociocchi, Advanced features in Bayesian reputation systems, in Trust, Privacy and Security in Digital Business, vol. 5695, Heidelberg: Springer, 2009, pp. 105-114.
[12] A. Jøsang, X. Luo, X. Chen, Continuous ratings in discrete bayesian reputation systems. In Proceedings of the Joint iTrust and PST Conferences on Privacy, Trust Management and Security (IFIPTM 2008), Trondheim, 2008.
[13] A. Jøsang and J. Haller, Dirichlet reputation systems, in Proceedings of the International Conference on Availability, Reliability and Security, Vienna, Austria, 2007.
[14] R. Jurca and B. Faltings, An incentive compatible reputation mechanism, In Proceedings of the IEEE Conference on E-Commerce, Melbourne, Australia 2003, pp.1026-1027.
[15] N. Littlestone and M. K. Warmuth, The weighted majority algorithm, Information and Computation, vol. 108, no. 2, pp. 212-261, 1994.
[16] R. Kerr, Smart cheaters do prosper: Defeating trust and reputation systems, in Proceedings of the 8th International Joint Conference on Autonomous Agents & Multiagent Systems, Budapest, Hungary 2009, pp. 999-1000.
[17] L. Mui, A. Halberstadt, and M. Mohtashemi, Notions of reputation in multi-agent systems: a review.In Proceedings of the First International Joint Conference on Autonomous Agents & Multiagent Systems, Bologna,Italy, 2002, pp 280-287.
[18] S. P. Marsh, Formalising Trust as a Computational Concept. PhD thesis, Department of Computing Science and Mathematics, University of Stirling, 1994.
[19] J. Patel, W. L. Teacy, N. R. Jennings, and M. Luck, A probabilistic trust model for handling inaccurate reputation sources, in Proceedings of 3rd International Conference on Trust Management, Rocquencourt, France, 2005, pp. 193-209.
[20] S. Ramchurn, D. Huynh, and N. R. Jennings, Trust in multi-agent systems, The Knowledge Engineering Review, vol. 19, no. 1, pp. 1-25, 2004.
[21] P. Resnick, R. Zeckhauser, J. Swanson, and K. Lockwood, The value of reputation on eBay: a controlled Experiment, in Proceedings of Esa Conference, Boston, Ma, 2002.
[22] S. Ruohomaa, L. Kutvonen, and E. Koutrouli, Reputation management survey, in Proceedings of the 2nd International Conference on Availability, Reliability and Security. Vienna, Australia. IEEE Computer Society, 2007, pp. 103-111.
[23] K. Regan, K. Cohen, and P. Poupart, The Advisor-POMDP: A principled approach to trust through reputation in electronic markets, in Proceedings of Conference on Privacy, Security & Trust, New Brunswich,Canada, 2005, pp. 121-130.
[24] J. Sabater, Evaluating the regret system, Applied Artificial Intelligence, vol. 18, no. 9-10, pp. 797-813, 2004.
[25] J. Sabater, and C. Sierra, REGRET: Reputation in gregarious societies, in Proceedings of the 5th International Conference on Autonomous Agents, Montreal,Quebec,Canada, 2001, pp. 194-195.
[26] J. Sabater, and C. Sierra, Social ReGret, A reputation model based on social relations. SIGecom Exchanges, vol. 3, no. 1, pp. 44-56, 2002.
[27] J. Sabater, and C. Sierra, Review on computational trust and reputation models, Artificial Intelligence Review, vol. 24, no. 1, pp. 33-60, 2004.
[28] G. Shafer, A Mathematical Theory of Evidence. NJ: Princeton University Press, 1976.
[29] W. T. L. Teacy, J. Patel, N. R. Jennings, and M. Luck, TRAVOS: Trust and reputation in the context of Inaccurate information sources, Autonomous Agents and Multi-Agent Systems, vol. 12, no. 2, pp. 183-198, 2006.
[30] Y. Wang and M. P. Singh,: Formal trust model for multiagent systems, in Proceedings of the 20th International Joint Conference on Artificial Intelligence, Hyderabad,India, 2007, pp. 1551-1556.
[31] A. Whitby, A. Jøsang, and J. Indulska, Filtering out unfair ratings in Bayesian reputation systems, in Proceedings of the 7th International Workshop on Trust in Agent Societies, New York, USA, 2004, pp 19-23.
[32] L. Xiong and L. Liu, A reputation-based trust model for peer-to-peer ecommerce communities, in Proceedings of IEEE Conference on E-Commerce, San Diego,CA,USA, 2003, pp 228-229.
[33] L. Xiong and L. Liu, PeerTrust: Supporting reputation-based trust for peer-to-peer electronic communities, IEEE Transactions on Knowledge and Data Engineering, vol. 16, no. 7, pp. 843-857, 2004.
[34] B. Yu and M. P. Singh, Detecting deception in reputation management, in Proceedings of the 2nd International Joint Conference on Autonomous Agents & Multiagent Systems, Melbourne, Australia, ACM, 2003, pp. 73-80.
[35] B. Yu and M. P. Singh , An evidential model of distributed reputation management, in Proceedings of the 1st International Joint Conference on Autonomous Agents & Multiagent Systems, Bologna, Italy, ACM, 2002, pp. 294-301.
[36] G. Zacharia and P. Maes, Trust management through reputation mechanisms. Applied Artificial Intelligence, vol. 14, no. 9, pp. 881-908, 2000.