A Dynamic Method for Mixing Direct and Indirect Trust Evidence

Han Yu
School of Computer Engineering
Nanyang Technological University, Singapore
yuhan@pmail.ntu.edu.sg

Zhiqi Shen
School of Computer Engineering
Nanyang Technological University, Singapore
zqshen@ntu.edu.sg

Chunyan Miao
School of Computer Engineering
Nanyang Technological University, Singapore
ascymiao@ntu.edu.sg

Cyril Leung
Department of Electrical and Computer Engineering
University of British Columbia, Canada
cleung@ece.ubc.ca

***Abstract*—As an increasing number of elderly people start to use online services, there is an urgent need to protect them from exploitation by malicious service providers (SPs). Testimonies and direct interaction experiences are both useful sources of trust evidence, but the weight given to each source to estimate a SP's trustworthiness is hard to determine. We propose a reinforcement learning based method to dynamically determine the optimal fix of direct and indirect trust evidence to help the elderly users form an accurate trust opinion on SPs. Extensive simulations have demonstrated the proposed method significantly outperforms existing approaches in terms of mitigating the adverse effect of unreliable third party testimonies.**

Keywords-trust; reputation; learning

I. INTRODUCTION

As online services grow in number and variety over the past decade, more and more people carry out transactions involving valuable resources (both tangible and intangible) with others through the Internet. Trust among these individuals is essential for transactions to happen. The trusting party is often exposed to risks caused by the behavior of the trusted party.

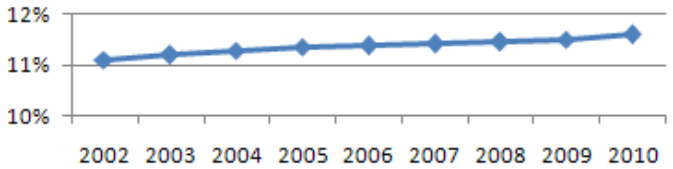

Figure 1. Worldwide age dependency ratio (data from the World Bank)

In recent years, the world has witnessed a growing global aging phenomenon. The age dependency ratio - the ratio of older dependents (people older than 64) to the working-age population (those ages 15-64) - has been consistently rising as shown in Fig. 1, which resulted in many elderly people in both the developed and the developing world having to live away from their children. This current cohort of the elderly populations is aging alongside the booming Internet. As they are expected to live independently for a longer time than their predecessors, it is reasonable to assume that the trend is towards the elderly populations increasingly lever online services to carry out more and more activities.

Trust is the extent to which one party is willing to depend on something or somebody in a given situation with a feeling of relative security, even though negative consequences are possible [1]. Without the mechanisms to gauge the interaction partner's social signals, an important part that help people judge the trustworthiness of a person in face to face interactions is missing. Therefore, computational trust models which evaluate one's trustworthiness through observing his past behaviors are used instead. Such models can be a valuable help to protect the elderly Internet users against malicious or selfish online service providers (SPs).

As online services flourish, so are mechanisms for people to provide feedback about their experience using them. Existing evidence based trust models often make use of two distinct sources of information to evaluate the trustworthiness of a service: 1) *direct trust evidence*: a service consumer's direct experience with a service provider, and 2) *indirect trust evidence*: the gossips about the service provider he receives from others. In these models, a service consumer uses past interaction outcomes with a service provider to estimate his future trustworthiness. Direct trust evidence is considered the most relevant to a service consumer. However, acquiring enough of it to make an accurate estimation of a service provider's trustworthiness requires large effort for the service consumer to explore and may expose him to a high level of risk of being cheated by unreliable service providers. Indirect trust evidence from witnesses (other consumer who interacted with a service provider before) can be useful in reducing a consumer's risk exposure. However, this introduces new uncertainties as the indirect trust evidence may be inaccurate due to malicious or selfish behavior by the witnesses. Thus, one of the most important research questions in evidence based trust models is how to mix the direct and indirect trust evidence about a service provider to accurately estimate its trustworthiness.

The majority of the existing trust models adopt a weighted average approach in mixing these two sources of evidence [2]. Direct trust evidence is often assigned a weight of γ ($0 \leq \gamma \leq 1$), and indirect evidence is then assigned a weight of $(1-\gamma)$. In most existing approaches, it is assumed γ is the same across all known service providers and its value is predetermined by domain experts [3]. In this paper, we propose a reinforcement learning based method to dynamically determine the weight assigned to each source of trust evidence based on the outcome of a consumer's interaction with service providers. The model learns a unique γ value for each service provider known by a

This research is supported, in part, by the Singapore Millennium Foundation (SMF) and the Singapore-MIT Alliance in Research and Technology (SMART) Grant No. ING11032-ICT.

consumer to enhance its accuracy. Experimental results demonstrate that the proposed method significantly outperforms existing approaches.

II. Related Work

Existing approaches for mixing direct and indirect trust evidence to evaluate a service provider's trustworthiness can be divided into two broad categories: 1) static approaches, where the value of γ is predefined, and 2) dynamic approaches, where the value of γ is updated during the lifetime of a service consumer.

A. Static Approaches

A long list of literature exists where static γ values are used in computational trust models. The majority of them tend to take a balanced approach by assigning the value of γ to 0.5 [4], [5], [6], [7]. In some more focused studies, the authors assigned the value 0 [8], [9], or 1 [10], [11] to γ to exclusively use only one source of trust information to make their decisions. Barber and Kim [12] have empirically proved, without considering the possibility of misleading indirect trust evidence, that direct experience is effective over the long term, but indirect trust evidence gives an accurate picture more quickly. Thus, by discarding one source or the other, approaches that assign 0 or 1 to γ forfeit part of the advantages of an evidence based trust model. However, using one static value for γ across all environmental conditions is not always an optimal approach.

B. Dynamic Approaches

Some existing work has explored varying the value of γ dynamically based on different assumptions. In [13], the authors alter the value of γ based on the number of direct observations on the behavior of a service provider available to a consumer. The paper assumes that every consumer starts from having no prior interaction experience with a service provider, and gradually accumulates direct trust evidence over time. It initially completely relies on indirect trust evidence to select service providers. As the number of repeated interactions with a service provider j increase, the value of γ also increases according to the formula $\gamma = \begin{cases} \frac{N_j^b}{N_{min}}, & if\ N_j^b < N_{min} \\ 1, & otherwise \end{cases}$, where N_j^b is the total number of direct observations of consumer b on service provider j, and N_{min} is the minimum number of direct observations determined by an acceptable level of error ε and a confidence c according to the *Chernoff Bound Theorem*: $N_{min} = -\frac{1}{2\epsilon^2}\ln(\frac{1-c}{2})$. This approach is not concerned with the accuracy of the indirect trust evidence when altering the value of γ. Rather, its aim is to accumulate enough direct trust evidence so that a consumer can make a statistically significant estimation on the trustworthiness of a service provider without relying on indirect trust evidence. In order to achieve a high level of confidence and low level of error, N_{min} can be very high. For example, a 90% confidence of at most 5% error requires at least 185 direct interactions. In real life applications, this may cause the service consumer to be exposed to significant risk of being cheated. In addition, since the value of γ eventually reaches 1, the approach is assuming that the behavior pattern of the service provider will never change. This may not always be true and it limits the applicability of the approach under certain circumstances.

Fullam and Barber [14] proposed a approach based the Q-learning technique [15] to select an appropriate γ value from a predetermined set of candidate values to be used to evaluation the trustworthiness of all potential service providers at each time step. The candidate γ values are set as $\{\gamma_1, \dots, \gamma_n\}$ assuming expert opinions about the underlying system characteristics are available. Based on the reward accumulated by a consumer under different γ values using the Q-learning technique at each time step, the γ value associated with the highest accumulated reward is selected. This work took the first step towards using real-time interaction outcomes to drive the consumer's subsequent interaction decisions. However, as the γ values are selected from a predetermined set, it is still partially relying on heuristics. The set of possible alternative γ values can impact the performance of the approach.

III. Problem Formulation

TABLE I. Notations used in this paper to describe the problem

c_i	A consumer
s_j	A service provider
w_k	A witness
$W_{i,j}$	A set of witnesses for s_j known to c_i.
$O_t^{i,j}$	The binary outcome of interaction between c_i and s_j at time t.
$d_t^{k,j}$	A testimony from w_k with regard to s_j at time t.
$D_{t,d}^{i,j}$	The binary decision by c_i on whether to interact s_j with at time t based on direct trust evidence only.
$D_{t,ind}^{i,j}$	The binary decision by c_i on whether to interact s_j with at time t based on indirect trust evidence only.
C	The cost in utility incurred by c_i when engaging the service of s_j.
G	The gain in utility after a successful interaction c_i and s_j.
$p_{m,t}$	The probability that testimonies from witnesses mislead the interaction decision of c_i with regard to s_j at time t.
$\gamma_{i,j}$	The weight assigned to direct trust evidence about s_j by c_i.

In this paper, we use the list of notations in Table I to describe the problem. We consider an open online system where the participants can play two types of roles:

1) *Service providers*, who provide services needed by others and do not need to rely on others to complete these services; and
2) *Consumers*, who need to rely on the service providers to accomplish certain goals.

In the eyes of a consumer, other consumers who provide it with indirect trust evidence (i.e., testimonies) about a service provider are regarded as *witnesses*.

Assuming at each time step t, a consumer c_i will at most interact with one s_j. Each time, c_i chooses the interaction partner from several candidate s_js according to their estimated trustworthiness. Whenever c_i needs to assess the trustworthiness of a service provider s_j, he draws upon both his own direct trust evidence with s_j (if there is any) as well as asks for testimonies from the set of witnesses $W_{i,j}$ who are known to have interacted with s_j before. A witness

w_k may reply c_i with a testimony $d_t^{k,j}$. A malicious w_k can be assumed to misbehave in two ways: 1) *consistent lying*, where w_k always replies with the opposite of what he believes the true trustworthiness of s_j should be following an independent and identically distributed (i.i.d.) lying probability; and 2) *collusion*, wherew_k only gives unfairly positive testimonies about a small set of s_js with whom they collude, but gives fair testimonies about others. All $d_t^{k,j}$ received by c_i are given equal weight when evaluating the trustworthiness of s_j. Depending on the presence of both types of the decision on which candidate s_j is to be selected for interaction by c_i at time step *t* might be adversely affected.

In order to hire s_j, c_i incurs a utility cost of *C*. If s_j successfully completes the task assigned to it by c_i, c_i receives a utility gain of *G*. In this paper, we assume $C < G$, and the values of *C* and *G* are the same for any pair of c_i and s_j. If c_i decides to interact with s_j at time step *t*, we assume the outcome of the interaction $O_t^{i,j}$ can be observed by c_i within the same time step. We assume that $O_t^{i,j}$ to be either successful $(O_t^{i,j} = 1)$ or unsuccessful $(O_t^{i,j} = 0)$. New w_k for s_j discovered by c_i over time are added into $W_{i,j}$. The objective is to maximize the time averaged utility gain of c_i throughout his lifetime in the presence of malicious s_js and witness agents

$$\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T}[(1-p_{m,t})\cdot(G-C)-p_{m,t}\cdot C] \qquad (1)$$

by adjusting the value of $\gamma_{i,j}$ based on $O_t^{i,j}$ over time through minimizing $p_{m,t}$. In an open system, the interactions between a consumer and a service provider do not have specific ending times. Therefore, the total number of interactions between any consumer and service provider pair cannot be known in advance.

IV. The Proposed Method

To solve the aforementioned problem, we propose a reinforcement learning (RL) based method to dynamically adjust the weight given to the two sources of trust evidence. The advantage of a RL based method is that it can learn directly from raw experience without a preconceived model of the environment dynamics. In addition, since RL does not need wait for the final outcome of a long series of interactions to be available before making estimations, it is suitable for our problem where interactions between any pair of s_j and c_i can go on indefinitely.

The proposed method is focused on addressing the problem of adapting a c_i's the reliance on the two sources of trust evidence during its lifetime. It does not matter which model c_i uses to evaluate the trustworthiness of a s_j. However, to facilitate the discussion in the following parts of this paper, we assume the popular Beta Reputation System (BRS) [16] is used as the underlying trust evaluation model for the consumers. The BRS estimates the trustworthiness of a s_j as follows:

$$\tau_{i,j} = E[\Pr(s_j)] = \frac{\alpha}{\alpha+\beta}$$

$$\alpha = N_p + 1, \beta = N_n + 1 \qquad (2)$$

where $\tau_{i,j}$ denotes the trustworthiness value of s_j based on the direct trust evidence from c_i. It is equivalent to the expectation of the probability that s_j will successfully serve c_i's requests in the future. N_p is the total number of successful interactions between s_j and c_i so far, and N_n is that of the successful interactions.

In the proposed method, c_i's actions regarding the choice between the two sources of trust evidence for any s_j is divided into two parts: 1) applying the current strategy (using the latest $\gamma_{i,j}$ value to estimate the trustworthiness of s_j, and 2) updating the $\gamma_{i,j}$ value based on the latest interaction outcome if s_j is chosen as the interaction partner for the last time step. These two parts as realized as the *actor module* and the *critic module* respectively according to the principles of the Actor-Critic learning method [15]. For each s_j known to a c_i, two critic module are present for the two sources of trust evidence and one actor module is present for estimating the trustworthiness of this s_j.

A. The Critic Module

The critic module in the proposed method performs the task of reward accumulation. Since the two critic modules for each known s_j are essentially the same but only taking in different sources of input data, in the following part, we will only discuss the critic module for direct trust evidence.

The value function of the critic module is designed based on the objective function presented in Eq. 1:

$$r_d = \frac{1}{T_{i,j}}\sum_{t=1}^{T_{i,j}}[(\mu_t\cdot(G-C)-(1-\mu_t)\cdot C)]$$

$$\mu_t = \begin{cases} 0, & if\ D_{t,d}^{i,j}=1 | O_t^{i,j}=0\ or\ D_{t,d}^{i,j}=0 | O_t^{i,j}=1 \\ 1, & if\ D_{t,d}^{i,j}=1 | O_t^{i,j}=1\ or\ D_{t,d}^{i,j}=0 | O_t^{i,j}=0 \end{cases}$$

$$D_{t,d}^{i,j} = \begin{cases} 0, & \tau_{i,j} < \theta \\ 1, & \tau_{i,j} \geq \theta \end{cases}. \qquad (3)$$

r_d can be considered as the discrete time averaged reward accumulated by the direct trust evidence source. $T_{i,j}$ represents the total number of interactions between s_j and c_i. μ_t is a variable determining whether the direct trust evidence source should be rewarded or punished at each time step. Its value toggles between 0 and 1 according to the relationship between the interaction decision $D_{t,d}^{i,j}$ as suggested by the direct trustworthiness evaluation $\tau_{i,j}$ and the observed actual interaction outcome $O_t^{i,j}$ after the overall interaction decision is made. As $D_{t,d}^{i,j}$ is only one component of the overall interaction,

it is possible that $D_{t,d}^{i,j}$ suggests not to interact with s_j and yet the overall decision is otherwise.

Once the latest r_d is available, it is compared with the baseline discrete time averaged reward accumulated by the direct trust evidence source $\widetilde{r_d}$ to update the learning parameter p_d according to:

$$p_d \leftarrow p_d + \rho \cdot (r_d - \widetilde{r_d}) \cdot (1 - \pi_d)$$

$$0 < \rho \leq 1. \quad (4)$$

ρ represent the rate of learning for the proposed method. It affects the size of change that can be achieved at each step of learning.

After p_d is updated, $\widetilde{r_d}$ is updated as follows:

$$\widetilde{r_d} \leftarrow \varphi \cdot \widetilde{r_d} + (1 - \varphi) \cdot r_d$$

$$0 < \varphi \leq 1. \quad (5)$$

$\widetilde{r_d}$ can be regarded as a basis for comparing whether c_i has become better off or worse off by aggregating the direct trust evidence into the estimation for the trustworthiness of s_j using the latest $\gamma_{i,j}$ value. φ is the weight assigned to the baseline value before it has been updated.

Similarly, the learning parameter p_{ind} for the indirect source of trust evidence can be obtained. When both p_d and p_{ind} are obtained, the learning parameters π_d and π_{ind} can be updated according to the Gibbs Softmax method [15] as:

$$\pi_d = \frac{e^{p_d}}{e^{p_d} + e^{p_{ind}}}$$

$$\pi_{ind} = \frac{e^{p_{ind}}}{e^{p_d} + e^{p_{ind}}}. \quad (6)$$

A desirable characteristic of this method is that π_d and π_{ind} can be regarded as the probability of selecting each source of trust evidence ($\pi_d + \pi_{ind} = 1$). Therefore, we assign $\gamma_{i,j}$ to be equal to π_d.

B. The Actor Module

The actor module is concerned with the strategy for aggregating the trust evidence from the two available sources. Since, in this paper, we do not keep track of the credibility of individual witnesses, a simple average over the available testimonies about s_j from all w_ks who have responded to c_i's request is taken. This value is then combined with $\tau_{i,j}$ using the following equation

$$rp_j = \gamma_{i,j} \cdot \tau_{i,j} + (1 - \gamma_{i,j}) \cdot \frac{1}{M} \sum_{k=1}^{M} d_t^{k,j}. \quad (7)$$

$d_t^{k,j}$ denotes a testimony about s_j fromw_k at time step *t*. It has the same format with $\tau_{i,j}$ since, assuming w_k is completely honest, he should have provided direct trustworthiness evaluation about s_j to c_i. rp_j represents the overall reputation of s_j which is used by c_i to estimate s_j's true trustworthiness. At each time step, c_i might have more than one candidate s_js to choose from. In this paper, we assume c_i will always select the s_j with the highest rp_j to interact with.

V. Experimental Evaluation

In order to evaluate the effectiveness of the proposed method against malicious witness populations, we designed a simulation test-bed with a simple economic system. Groups of autonomous consumer agents equipped with three static methods and two state-of-the-art dynamic methods for determining the value of $\gamma_{i,j}$ compete with those equipped with the proposed method in the test-bed. Their performances are judged with practical metrics. Under two different sets of experiment conditions, it has been observed that the proposed method significantly outperforms all other methods in terms of reducing the normalized average utility loss of the consumers.

A. Experiment Environment

The environment of each experiment consists of $N_s = 100$ service provider agents with different behavior patterns and $N_w = 100$ common witness agents who accumulate direct trust evidence about service provider agents and provide testimonies to requesting consumer agents following different strategies. These agents are allowed to run for T_s time steps to accumulate some direct trust evidence before agents equipped with various $\gamma_{i,j}$ calculation methods join the environment. Since it is not possible to enumerate all service provider population behavior patterns, we select one which consists of 20% honest SPs (agents who complete their assigned tasks successfully 90% of the time), 40% unreliable SPs (agents who complete their assigned tasks successfully 30% of the time), and 40% malicious SPs (agents who complete their assigned tasks successfully 10% of the time). This is a highly hostile SP population. It is assumed that if the proposed method can perform well under such a hostile environment, it should perform well under more accommodating environments, too. Thus, in the experiments, we only need to alter the behavior pattern of the witness agents.

The common witness agents follow three behavior patterns:

1) *Honest*, honest witness agents give out their untempered direct trustworthiness evaluations as testimonies 90% of the time;
2) *Badmouthing (BM)*, BM witness agents give out unfairly negative testimonies 90% of the time;
3) *Ballot-stuffing (BS)*, BS witness agents give out unfairly positive testimonies 90% of the time.

In each experiment, the composition of the common witness agent population is altered to simulate different conditions. In the following sections, *Hon* denotes a population consisting entirely of honest witness agents. *BMn* denotes a population consisting of *n*% BM witness agents and the rest are honest witness agents. *BSn* denotes a population consisting of *n*% BS witness agents and the rest are honest witness agents.

Five groups of consumer agents are used to compete with the group of consumer agents equipped with the proposed method. They are:

1) *Group* $\gamma = 0$, agents who completely rely on indirect trust evidence;

2) *Group* $\gamma = 0.5$, agents who rely on a balanced fix of direct and indirect trust evidence;

3) *Group* $\gamma = 1$, agents who completely rely on direct trust evidence;

4) *Group M2002*, agents who use the method described in [13] to set the γ value;

5) *Group F&B2007*, agents who use the method described in [12] to set the γ value.

The group of agents equipped with the proposed method is labeled as Group *Y2012*. Each group consists of *N*=10 agents and they only requests for testimonies from the common witness agent group. Each consumer agent needs to complete 200 tasks with SPs. The values selected by the parameters involved in the proposed method are listed in Table II.

TABLE II. PARAMETER VALUES USED IN THE SIMULATION

φ	ρ	C	G	θ	M
0.6	0.4	1	5	0.5	5

B. Evaluation Metric

The normalized average utility gain $0 \leq \sigma \leq 1$ is used to gauge the performance of various groups of competing agents in the experiments. It is calculated as

$$\sigma = \frac{\frac{1}{T \cdot N}\sum_{t=1}^{T}\sum_{i=1}^{N} g_{i,t} - C}{G - C}. \quad (8)$$

T is the total number tasks completed by each c_i, N is the number of consumer agents in each group, and $g_{i,t}$ is the actual utility gain of each c_i at time step t.

C. Observations

The first set of experiments is conducted under non-collusive common witness agent populations. The common witness population composition is altered from *BM100* to *Hon* and then to *BS100* to test the performance of different groups of consumer agents. The results are summarized in Fig. 2 and Fig. 3. It can be observed that Group $\gamma = 1$ achieved the lowest σ value as they need more exploration to identify the trustworthy SPs. Completely relying on indirect trust evidence is also not a good strategy as the performance of Group $\gamma = 0$ is heavily affected by the presence of unreliable witness agents of both BM and BS types. However, the saving in exploration resulted from completely relying on testimonies from other did allow Group $\gamma = 0$ to achieve higher σ values than Group $\gamma = 1$ except under extremely hostile witness agent populations of *BM100* and *BS100*. The performance of the Group $\gamma = 0.5$ is the best among the three groups using static γ values. Group *F&B2007*'s performance is similar to that of Group *M2002*. As *F&B2007* tries to learn which static strategy $(\gamma = 0, 0.5, or\ 1)$ is the best under different conditions, its performance more or less tracks that of Group $\gamma = 0.5$. Group *Y2012* outperforms all other methods under all testing conditions.

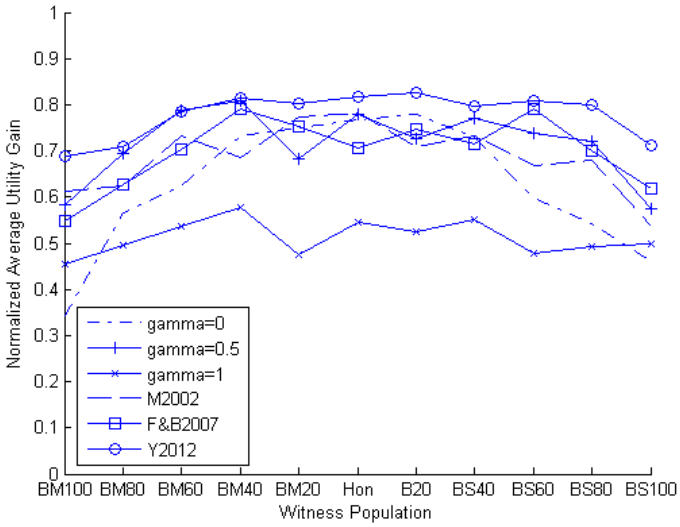

Figure 2. The performance of various methods against different non-colluding witness population configurations

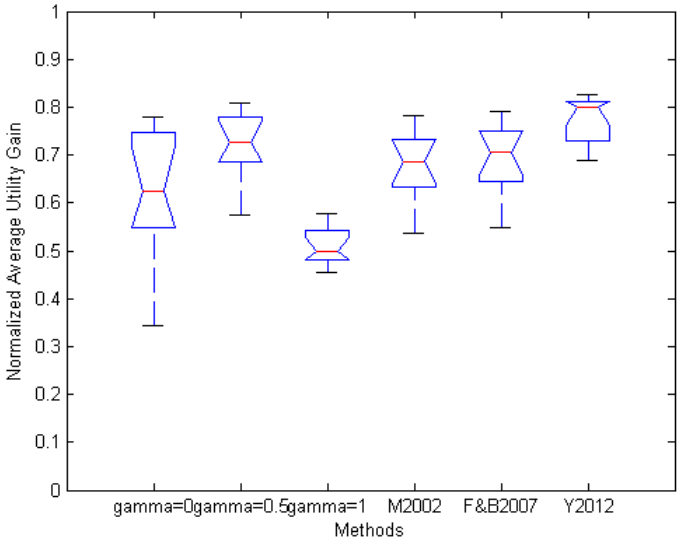

Figure 3. Boxplots comparing the performance of various methods against different non-colluding witness population configurations

The second set of experiments is conducted under collusive common witness agent populations. Under this condition, a BS witness agent only gives out unfairly positive testimonies when the SP of interest is a malicious SP. The common witness population composition is altered from *Hon* to *BS100*. The results are summarized in Fig. 4 and Fig. 5. The relative performances of different groups are similar to non-collusive conditions. The improvement of Group *Y2012* in terms of average percentage reduction in normalized average utility loss $(1 - \sigma)$ is significant against all other groups as shown in Table III. Nevertheless, the improvements under collusive conditions are lower than under non-collusive conditions. This is due to the fact that collusive witness agents appear to behavior strategically (provide reliable testimonies about non-colluding SPs, but provide unfair testimonies about colluding SPs). Without evaluating the credibility of individual witness agents, collusion is difficult to defend against.

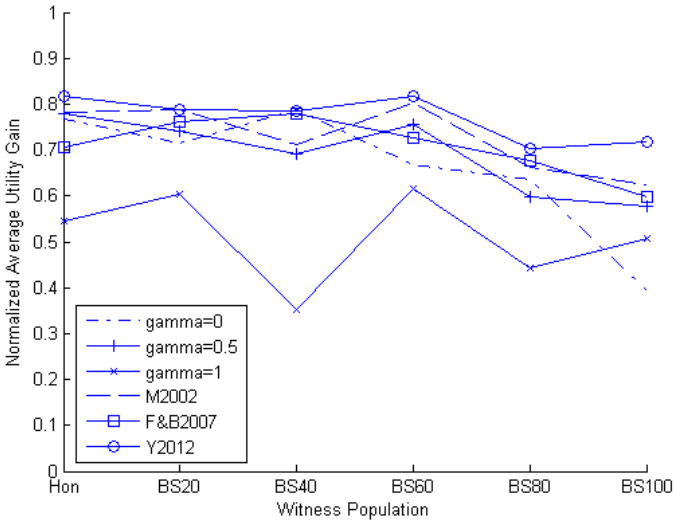

Figure 4. The performance of various methods against different colluding witness population configurations

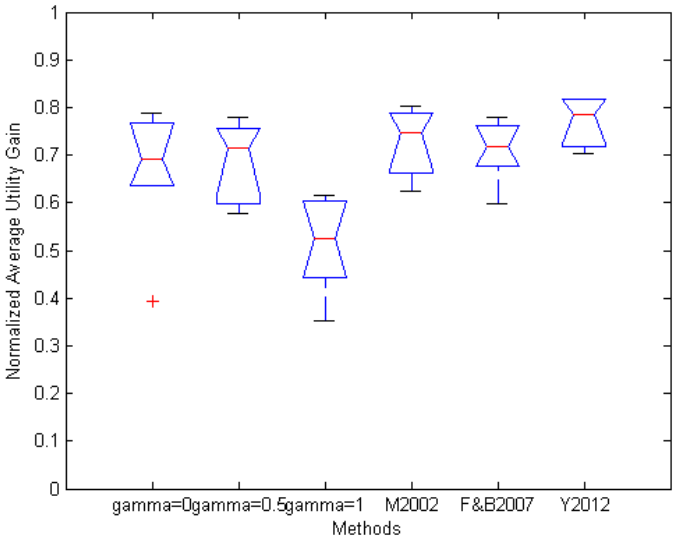

Figure 5. Boxplots comparing the performance of various methods against different colluding witness population configurations

TABLE III. IMPROVEMENT OVER EXISTING METHODS (% REDUCTION IN NORMALIZED AVERAGE UTILITY LOSS)

Methods		**Malicious Witness Population Behavior**	
		Non-Collusion	**Collusion**
Static Methods	**$\gamma = 0$**	40.48%	32.42%
	$\gamma = 0.5$	22.09%	26.04%
	$\gamma = 1$	54.48%	53.23%
Dynamic Methods	**M2002**	29.46%	15.58%
	F&B2007	26.01%	21.66%

VI. CONCLUSIONS AND FUTURE WORK

This paper proposed a method to dynamically learn the optimal mix of direct and indirect trust evidence for estimating the trustworthiness of SPs. The method has significantly outperformed existing approaches under extensive simulations and has the potential to help elderly users select trustworthy online service providers under highly hostile witness populations. In subsequent research, we will incorporate learning based witness credibility evaluation mechanisms to improve the effectiveness of the proposed method against collusion.

REFERENCES

[1] A. Jøsang, R. Ismail, and C. Boyd, "A survey of trust and reputation systems for online service provision," Decision Support System, vol. 43(2), pp.618-644, 2007.

[2] Z. Noorian and M. Ulieru, "The state of the art in trust and reputation systems: a framework for comparison," Journal of Theoretical and Applied Electronic Commerce Research, vol. 5(2), pp.97-117, 2010.

[3] H. Yu, Z. Shen, C. Miao, C. Leung, and D. Niyato, "A survey of trust and reputation management systems in wireless communications," Proceedings of the IEEE, vol. 98(10), pp.1755-1772, 2010.

[4] J. Weng, C. Miao, and A. Goh, "An entropy-based approach to protecting rating systems from unfair testimonies," IEICE Transactions, vol. 89-D(9), pp.2502-2511, 2006.

[5] J. Weng, Z. Shen, C. Miao, A. Goh, and C. Leung, "Credibility: How Agents Can Handle Unfair Third-party Testimonies in Computational Trust Models," IEEE TKDE, vol.22(9), pp.1286-1298, 2010.

[6] S. Liu, J. Zhang, C. Miao, Y-L. Theng, and A. C. Kot, "iCLUB: an integrated clustering-based approach to improve the robustness of reputation systems," In Proc. AAMAS-11, pp.1151-1152, 2011.

[7] H. Yu, S. Liu, A. Kot, C. Miao, and C. Leung, "Dynamic witness selection for trustworthy distributed cooperative sensing in cognitive radio networks," In Proc. of the 13th IEEE ICCT, pp.1-6, 2011.

[8] C.M. Jonker and J. Treur, "Formal analsys of models for the dynamics of trust based on experiences," In Proc. of MAAMAW-99, pp.221-231, 1999.

[9] M. Schillo, P. Funk, and M. Rovatsos, "Using trust for detecting deceitful agents in artificial societies," In Proc. of the Applied Artificial Intelligence Journal, pp.825-848, 2000.

[10] J. Shi, G. Bochmann, and C. Adams, "Dealing with recommendations in a statistical trust model," In Proc. of the Trust in Agent Societies Workshop at AAMAS-05, pp.144-155, 2005.

[11] L. Teacy, J. Patel, N.R. Jennings, and M. Luck, "Coping with inaccurate reputation sources: experimental analysis of a probabilistic trust model," In Proc. of AAMAS-05, pp.997-1004, 2005.

[12] K.S. Barber and J. Kim, "Soft security: isolating unreliable agents from society," Trust, Reputation and Security: Theories and Practice, Springer, pp.224-233, 2003.

[13] L. Mui, M. Mohtashemi, and A. Halberstadt, "A computational model of trust and reputation," In: Proc. of HICSS, 2002.

[14] K.K. Fullam and K.S. Barber, "Dynamically learning sources of trust information: experience vs. reputation," In Proc. of AAMAS-07, 2007.

[15] R. S. Sutton and A. G. Barto, "Reinforcement learning: an introduction," MIT Press, 342p., 1998.

[16] A. Jøsang and R. Ismail, "The beta reputation system," In Proc. of the 15th Bled Conference on Electronic Commerce, 2002.