Combining Direct Trust and Indirect Trust in Multi-Agent Systems

Elham Parhizkar[1] , **Mohammad Hossein Nikravan**[1] , **Robert C. Holte**[2] and **Sandra Zilles**[1]

[1]Department of Computer Science, University of Regina, Canada
[2]Department of Computing Science, University of Alberta, Canada
parhizkar@uregina.ca, nikravam@uregina.ca, rholte@ualberta.ca, zilles@cs.uregina.ca

Abstract

To assess the trustworthiness of an agent in a multi-agent system, one often combines two types of trust information: *direct trust* information derived from one's own interactions with that agent, and *indirect trust* information based on advice from other agents. This paper provides the first systematic study on when it is beneficial to combine these two types of trust as opposed to relying on only one of them. Our large-scale experimental study shows that strong methods for computing indirect trust make direct trust redundant in a surprisingly wide variety of scenarios. Further, a new method for the combination of the two trust types is proposed that, in the remaining scenarios, outperforms the ones known from the literature.

1 Introduction

Trust and reputation systems constitute an active branch of research in multi-agent systems. In various application domains, agents interact with one another in order to collect information, goods, or services that help with completing a set task. For such interactions to be largely successful, agents try to estimate how trustworthy other individual agents are.

While the approaches to modeling trust cover a wide variety of techniques, usually the literature distinguishes between methods for computing direct trust and those for computing indirect trust [Jøsang *et al.*, 2007]. Direct trust of agent A in agent B refers to the trust A builds in B solely based on past interactions between A and B. By contrast, agent A's indirect trust in B is based on recommendations (e.g., ratings) about B that one or more third-party advisors have provided to A.

Indirect trust can be difficult to establish since the third-party advisors themselves could be unreliable, for various reasons. Direct trust information can be unreliable as well, the most common reason being that agent A simply doesn't have enough experience of directly interacting with agent B to assess B's trustworthiness with confidence. Intuitively, agent A should combine indirect and direct trust information on agent B in such a way that, over time, the role of direct trust grows with the number of interactions A has had with B.

The questions addressed in our study are: How to best combine direct and indirect trust information? Are there typical situations in which one should ignore direct trust?

To answer these questions, we designed a broad empirical study, including (i) four methods for computing indirect trust known from the literature, namely ITEA [Parhizkar *et al.*, 2019], TRAVOS [Teacy *et al.*, 2006], ACT [Yu *et al.*, 2014], and MET [Jiang *et al.*, 2013], (ii) 14 different settings of dishonest third-party advisors as attacks against these systems, and (iii) five different methods for combining direct and indirect trust, three of which are novel. For computing direct trust itself, we focused on the Beta Reputation System [Jøsang and Ismail, 2002].

Our analysis reveals a number of interesting insights. First, it is rarely useful to rely only on direct trust; indirect trust information almost always helps. Second, in many scenarios, even facing an advisor pool in which 80% of the advisors are dishonest, the best indirect trust methods do not benefit from incorporating direct trust in decision-making at all. In several cases, it is even hurtful to incorporate direct trust. Third, direct trust tends to be more beneficial in cases when (i) all advisors are dishonest, or (ii) many advisors change their behavior dynamically. Case (ii) applies to many attack scenarios studied in the literature, so that direct trust cannot be ignored in general. Fourth, a core insight is that, among all the methods tested for combining direct and indirect trust, there is a clear winner: We propose a method of exponential decay of the weight of indirect trust over time. With very few exceptions, this new method either outperforms or is on par with all other tested methods for combining direct and indirect trust. Moreover, combining this method with ITEA, one obtains a system that is highly robust across the wide variety of scenarios we tested.

2 Preliminaries and Related Work

As in many studies in the literature, we assume that interactions with a trustee (the agent with which to interact) have binary outcomes, i.e., they can be either positive or negative. The relative frequency of positive interactions with a trustee can then be seen as that trustee's trustworthiness.

The vast majority of trust systems compute direct trust information using the Beta Reputation System (BRS) [Jøsang and Ismail, 2002]. Following the notation in [Parhizkar *et al.*, 2019], trustees are denoted s_j, indexed by j. Then the direct

trust an agent has in trustee s_j is given by

$$\text{brs}(p_j, n_j) = \frac{p_j + 1}{p_j + n_j + 2},\qquad(1)$$

where p_j and n_j refer to the number of positive, resp. negative, interactions that the agent has had with s_j. A larger value of this measure suggests a higher trustworthiness of s_j, but many systems store the numbers p_j and n_j separately, since two pairs (p, n) and (p', n') may result in the same value in brs, yet $p + n$ could be much larger than $p' + n'$, suggesting a larger confidence in the trustworthiness value derived from (p, n) than in that derived from (p', n'), since the former stems from a larger number of interactions.

By comparison, there is much more variety in the literature on indirect trust [Teacy *et al.*, 2006; Regan *et al.*, 2006; Jiang *et al.*, 2013; Yu *et al.*, 2014; Teacy *et al.*, 2012; Yu and Singh, 2003; Irissappane and Zhang, 2017; Liu *et al.*, 2017; Liu *et al.*, 2011; Cohen *et al.*, 2018; Weng *et al.*, 2010; Parhizkar *et al.*, 2019; Jøsang and Ismail, 2002]. Some systems identify unreliable recommendations by verifying that recommendations of an individual advisor display substantially different statistical properties than those of the whole set of advisors [Jøsang and Ismail, 2002]. Others make a truster compare its own direct experience to the recommendations of an advisor [Teacy *et al.*, 2006; Yu and Singh, 2003]. Once detected, unreliable recommendations are often removed [Jøsang and Ismail, 2002; Jiang *et al.*, 2013; Yu *et al.*, 2014] or down-weighted [Teacy *et al.*, 2006]. Some probabilistic systems [Regan *et al.*, 2006; Teacy *et al.*, 2012] instead exploit correlations between direct experience and advisors' recommendations in order to implicitly modify (re-interpret) recommendations that are likely dishonest.

In our study, we compare four systems for indirect trust, namely ITEA [Parhizkar *et al.*, 2019], TRAVOS [Teacy *et al.*, 2006], ACT [Yu *et al.*, 2014], and MET [Jiang *et al.*, 2013]. ITEA, TRAVOS, and MET were chosen because a recent study identified these three methods as state-of-the-art for indirect trust and they are often competitive with each other [Parhizkar *et al.*, 2019]. While the same study ranked ACT somewhat inferior in terms of indirect trust, we include ACT in the present analysis, since it is, to the best of our knowledge, the only existing method using reinforcement learning for combining direct and indirect trust. Of the other three methods, only TRAVOS combines direct and indirect trust. Our study will compare TRAVOS's and ACT's approaches for combining direct and indirect trust to several others we propose, some of which are simple baseline methods. It will turn out that one of our newly proposed methods outperforms all existing combination methods, independent of the tested indirect trust method.

3 Methods for Combining Trust Sources

A straightforward approach to combining direct and indirect trust would be to assign each a fixed weight and calculate the weighted average. Intuitively though, one would want to begin with a low weight for direct trust and gradually, as more direct experience is collected, increase the weight of direct trust. TRAVOS ignores indirect trust once its direct trust is

considered strong enough; before that it integrates direct trust with the advisors with a fixed weight of $1/(k + 1)$, where k is the number of advisors and each advisor also has weight $1/(k + 1)$. ACT instead updates weights for combining the two trust sources after each iteration, using a reinforcement learning method. In our experiments, we incorporate ACT's combination method also in other systems, to see how it performs in comparison to other combination approaches. The combination methods we tested are the following:

Only Indirect. This method only uses indirect trust; direct trust is ignored completely. This is a baseline to test if combining direct and indirect trust actually leads to improvement.

Exponential Function (ExpFun). The weight of the indirect trust value, $w_{ind}(j)$, when the truster estimates the trustworthiness of s_j, is calculated as follows:

$$w_{ind}(j) = \exp^{-\lambda(p_j + n_j)},\qquad(2)$$

where $p_j + n_j$ is the number of interactions the truster has had with s_j. A similar formula was used in [Huynh *et al.*, 2006] in order to decrease the weight of individual ratings (advisor recommendations) over time, so as to make a trust system responsive to the recent changes in the behavior of individual agents. However, to the best of our knowledge, this function has not been used for combining direct and indirect trust before. The value of λ is set to 0.1 for all experiments.

ACT-RL. To derive the weights of direct and indirect trust values, ACT's reinforcement learning technique as described in [Yu *et al.*, 2014] is used. The initial weight is 0.5.

Average. Trustworthiness values are calculated as the average of direct and indirect trust values.

Direct as Advisor. This method uses a chosen indirect trust method and treats direct trust as if it were an additional advisor. In other words, given k advisors, direct trust is used as a $(k + 1)$-st advisor, and it is not distinguished from other advisors, except that it is always honest.

Online Learning. Here, the weights of direct and indirect trust are calculated with the ITEA algorithm [Parhizkar *et al.*, 2019] as if these two sources of trust were two experts.

TRAVOS+. TRAVOS has its own method for combining direct and indirect trust [Teacy *et al.*, 2006]; we tried it only within TRAVOS itself.

Exponential Function, Online Learning, and Direct as Advisor – though simple methods – were, to the best of our knowledge, never suggested for combining direct and indirect trust in the literature. Proposing and testing these methods is one of the contributions of our work.

4 Experimental Setup

We ran each of ITEA, TRAVOS, ACT, and MET in combination with every method in Section 3. The performance of each combination was measured in terms of relative frequency of unsuccessful interactions (RFU) [Parhizkar *et al.*, 2019; Yu *et al.*, 2014]. RFU is simply the ratio of the number of negative interactions over the number of all interactions. We

measure RFU for a specific target number of positive interactions (in our experiments 500).[1]

All interactions are independent random events. Each trustee has a constant trustworthiness value in $[0, 1]$ corresponding to the probability of a positive outcome when interacting with that trustee. Our empirical study uses one truster, 10 trustees, and 10 advisors.[2] As in [Parhizkar *et al.*, 2019], possible trustworthiness values of the trustees are 0.1, 0.2, ..., 0.9. Every number reported in our tables is an average of 100 runs of the same setting; each run has a preprocessing phase in which (i) the trustworthiness value of each trustee is initially sampled uniformly at random, and (ii) 1000 advisor/trustee pairs (a_i, s_j) are sampled at random. Each time a pair (a_i, s_j) is sampled, an interaction between a_i and s_j is simulated, using the trustworthiness value previously sampled for s_j. The outcome of the interaction is recorded by a_i.

Honest advisors will later always report $\text{brs}(p_j, n_j)$ (when using ITEA, ACT, or MET) or (p_j, n_j) (when using TRAVOS, which needs to handle pairs explicitly) as a recommendation on s_j, where p_j and n_j refer to the number of positive, resp. negative, interactions between the advisor and s_j during preprocessing.

4.1 Types of Dishonest Advisors

Our evaluation considers 14 settings, which differ in how the advisors distort the actual values of p_j and n_j for the trustees. The first 10 settings are identical to those in [Parhizkar *et al.*, 2019], just ordered and numbered slightly differently. Settings 11-14 are designed to prevent a truster from making a good choice based on the advisors' inputs only. This is done by making the behavior of advisors dynamic, i.e., advisors switch from a consistently honest behavior to a consistently dishonest behavior or vice versa, potentially multiple times. When an advisor is being dishonest in any of these last four settings, it returns (n_j, p_j) for trustee s_j instead of (p_j, n_j).

Setting 1: Partly Random Advisors. In preprocessing, a partly random advisor picks trustees for which it will be dishonest. Each trustee has a 50% chance of being picked. For each trustee s_j that is picked, the advisor randomly selects some $z \in (0, 1)$, computes any pair (p'_j, n'_j) of non-negative integers with $\text{brs}(p'_j, n'_j) = z$, and always reports (p'_j, n'_j) about s_j. For the remaining trustees, the advisor is honest.

Setting 2: Fully Random Advisors. These are partly random advisors (Setting 1) that are dishonest about all trustees.

Settings 3, 4: Badmouthing (BM)/Ballot-Stuffing (BS) Advisors. In the preprocessing phase, a BM/BS advisor picks trustees for which it will be dishonest. Each trustee has a 50% chance of being picked. For trustees that are not picked, the advisor is honest. For each trustee s_j that is picked, a BM advisor returns the (p, n) pair recorded in its preprocessing phase for which the $\text{brs}(p, n)$ value was smallest over all

trustees. A BS advisor returns the (p, n) pair of the trustee with the highest $\text{brs}(p, n)$ value recorded in preprocessing.

Settings 5, 6: Additive BM/BS Advisors. This type of advisor is adapted from [Yu *et al.*, 2014]. An additive BM advisor does the following for each trustee s_j: it first samples a random number $z \in [0.8, 1]$ and computes $z^* = \frac{p_j+1}{p_j+n_j+2} - z$. If $z^* > 0$, a pair (p, n) with $\text{brs}(p, n) = z^*$ is returned; otherwise $(0, p_j + n_j)$ is returned. An additive BS advisor uses $z^* = \frac{p_j+1}{p_j+n_j+2} + z$. If $z^* < 1$, a pair (p, n) with $\text{brs}(p, n) = z^*$ is returned; otherwise $(p_j + n_j, 0)$ is returned.

Settings 7, 8: All-Negative/All-Positive Advisors. For each trustee, an all-negative advisor reports $p = 0$ and $n = 1,000,000$, and an all-positive advisor reports $p = 1,000,000$ and $n = 0$.

Settings 9, 10: Selective BM/BS Advisors. For trustee s_j a selective BM advisor reports $(0, p_j + n_j)$ if $\text{brs}(p_j, n_j) \geq 0.5$ and is honest otherwise. A selective BS advisor reports $(p_j + n_j, 0)$ if $\text{brs}(p_j, n_j) \leq 0.5$ and is honest otherwise.

Setting 11: Camouflage. These advisors behave honestly for 100 interactions to develop their trustworthiness and then start to behave dishonestly.

Setting 12: Good-Bad-Good. These advisors behave honestly for 100 interactions, dishonestly for the next 100 interactions, and then honestly for all remaining interactions.

Setting 13: Periodic Behavior. These advisors alternate between honest and dishonest behavior periodically. They act honestly for the first 10 interactions, dishonestly for the next 10 interactions, honestly for the next 10 interactions, etc.

Setting 14: Random Behavior. On each iteration, these advisors choose between honest and dishonest behavior randomly with equal probability.

4.2 Further Details

The truster computes its direct trust in a trustee s_j using BRS, based on its direct interactions with s_j, starting with an empty history. Each experiment then proceeds in rounds until the truster has accumulated a target number of positive interactions (here 500) in total.

Suppose an indirect trust system IT and a method M for combining direct and indirect trust are chosen (when only indirect trust is used, the combination method is to ignore direct trust). Then, each round proceeds as follows. (i) The truster collects indirect trust information through IT and combines it with its direct trust information via M, to obtain trustworthiness values for all trustees. (ii) The truster interacts with the trustee whose trustworthiness value is maximal (breaking ties randomly). Depending on the outcome, the truster updates its direct trust information using BRS and its indirect trust information using IT. If the target number of positive interactions is not yet achieved, then the next round is initiated.

We set all parameters for all the tested systems exactly as in [Parhizkar *et al.*, 2019], except for the number of advisors in MET's trust networks. This number is usually chosen to be a constant that is at least half the total number of advisors. Since our simulations used $k = 10$ advisors, we set $n = 6$ as the trust network size for MET.

---

[1]This stands in contrast to measures that capture how well an agent estimates the trustworthiness of each individual trustee.

[2]Most empirical studies on indirect trust use a larger number of advisors (e.g., 100 in [Parhizkar *et al.*, 2019], 40 in [Jiang *et al.*, 2013]), but we decreased this number to give direct trust a better chance of contributing.

5 Results and Discussion

For each setting, we measured RFU after 500 positive interactions for four different percentages (20%, 40%, 60%, 80%) of dishonest advisors chosen at random from the set of all advisors. Tables 1–3 contain subtables of results for all the settings except Settings 5/6, which showed similar trends as Settings 3/4 and were omitted due to space constraints. Each table has a section for each system tested (ITEA, TRAVOS, ACT, MET), containing a row for each combination method from Section 3. Rows marked ACT/AC refer to the full ACT system. TRAVOS+ rows refer to the full TRAVOS system.

A bold entry indicates a statistically significant difference (using a two-tailed paired t-test at the 95% confidence level) compared to that system using indirect trust alone (the top row in each section). An asterisk in addition means a worse RFU compared to when using indirect trust alone.

Note that a system using direct trust alone is in most cases inferior to a system combining direct and indirect trust. The RFU value for a system using only direct trust in our settings would be 0.232 (since such a system is indifferent to the advisor setting, there is just one single number to report for it). This number is worse than all numbers we obtained for Settings 1 and 3–6, across all the systems/combinations. It is also worse than all numbers we report for TRAVOS, except for the case of 80% Selective BM Advisors (Table 2c), which is in our study by far the most difficult setting for establishing good indirect trust. ITEA with any combination method is also rarely worse than only direct trust (and usually substantially better); on top of the case of 80% Selective BM Advisors, it loses to direct trust just in a few cases in the Camouflage setting (Table 3a) and for Periodic Behavior (Table 3c). For ACT and MET, inferiority to only direct trust happens in a larger number of cases. However, ACT and MET appear to be overall the weaker indirect trust methods in our study; they are not competitive with ITEA or TRAVOS.

The results for MET paint a simple picture. With one exception (Selective BM (Table 2c), 80%), MET's RFU for Only Indirect is substantially worse than ITEA's. Hence, not surprisingly, direct trust leads to significant improvements to MET's Only Indirect RFU much more often than it does for ITEA's. However, in many settings, none of the methods of combining direct and indirect trust gives MET a better RFU than ITEA's RFU for Only Indirect. In other words, the use of direct trust does not make up for the deficiencies in how MET uses indirect trust. The settings in which this is not true are: Selective BM (Table 2c), Camouflage (Table 3a), Good-Bad-Good (Table 3b), Periodic Behavior (Table 3c), and Random Behavior (Table 3d); here direct trust significantly improves the indirect trust RFU for both ITEA and MET, and ITEA emerges the winner except for Selective BM.

The overall picture for ACT is virtually identical to that of MET. Again, with just one exception (Selective BM (Table 2c), 40%), ACT's RFU for Only Indirect is substantially worse than ITEA's. Direct trust leads to significant improvements to ACT's RFU much more often than it does for ITEA's. Also as with MET, in many settings, no combination with direct trust results in ACT having a better RFU than ITEA's RFU for Only Indirect (there are a few small, isolated

exceptions). The exceptions, for ACT, are: Fully Random Advisors (Table 1b), Selective BM (Table 2c), Camouflage (Table 3a), Good-Bad-Good (Table 3b), and Periodic Behavior (Table 3c).

The relation between TRAVOS and ITEA is not as straightforward as that of MET and ACT. There are still several settings where TRAVOS's best RFU score does not exceed ITEA's Only Indirect RFU. Then there are settings for which TRAVOS's best RFU is the same or better than ITEA's Only Indirect RFU, but not better than ITEA's best RFU. There is one setting that does not fall into either of the preceding categories: Selective BM advisors (Table 2c). In this setting, ITEA's Only Indirect RFU is better than TRAVOS's best RFU in the 20% and 40% columns, but TRAVOS's best RFU is better than ITEA's best RFU in the 60% and 80% columns.

Based on this analysis, ITEA outperforms TRAVOS, and both are superior to MET and ACT. As for when to combine direct and indirect trust, it becomes evident that ITEA and TRAVOS benefit less frequently from combining with direct trust than MET and ACT do. A reason for this could be that ITEA and TRAVOS are often more effective at computing indirect trust than ACT and MET. By the time direct trust becomes reliable, presumably the indirect trust components of ITEA and TRAVOS are as reliable as direct trust, making the latter redundant. However, there are general types of situations in which ITEA and/or TRAVOS benefit greatly from incorporating direct trust. Camouflage, Good-Bad-Good, and Periodic Behavior, i.e., settings in which advisors display a certain behavior consistently for an extended period of time, but then change that behavior, are hard to handle for indirect trust methods; here incorporating direct trust is helpful for ITEA and TRAVOS. Not reported in our tables are also experiments with 100% dishonest advisors, for which, not surprisingly, all indirect trust methods, across most advisor settings, improved upon incorporating direct trust.

5.1 Comparison of Combination Methods

Table 4 has a row for each setting and a column for each method of combining direct and indirect trust. Each cell in the table has two numbers. The first number, d, is how many times, in that setting, the particular combination method led to a significant decrease in RFU: in Tables 1–3 these are the bold numbers without an asterisk. The second number, i, is how many times, in that setting, the particular combination method led to a significant increase in RFU: in Tables 1–3 these are the numbers with an asterisk.

In each row, bold font indicates the entries in that row that dominate the others. (d_1, i_1) dominates (d_2, i_2) if either (i) $d_1 > d_2$ and $i_1 \leq i_2$, or (ii) $i_1 < i_2$ and $d_1 \geq d_2$. In some rows there is not a unique method that dominates.

Overall, our newly proposed method ExpFun is the best strategy, by a large margin, for combining direct and indirect trust, although we never attempted to optimize the parameter λ in Equation (2). Based on this analysis, in these settings, none of the other combination methods is worth considering.

5.2 Comparison of Combined Systems

We call a trust system A preferable to a trust system B in setting S if A's RFU value in S for $q\%$ dishonest advisors is

	20%	40%	60%	80%	20%	40%	60%	80%	20%	40%	60%	80%	20%	40%	60%	80%
ITEA/OI	0.142	0.146	0.152	0.152	0.147	0.166	0.170	0.202	0.147	0.159	0.159	0.170	0.140	0.141	0.141	0.144
ITEA/EF	0.144	0.144	0.149	0.152	0.153	0.159	**0.158**	**0.182**	0.145	0.152	0.155	0.165	**0.149***	0.147	0.144	0.142
ITEA/AC	**0.154***	0.152	0.154	**0.161***	**0.155***	0.166	0.170	0.195	0.155	0.164	0.158	0.176	**0.146***	**0.154***	**0.152***	**0.153***
ITEA/OL	**0.154***	**0.153***	0.153	**0.162***	**0.155***	0.167	0.169	0.196	0.153	0.164	0.157	0.179	**0.147***	**0.152***	**0.152***	**0.153***
ITEA/Avg	**0.154***	0.153	0.153	**0.162***	**0.155***	0.164	0.169	0.196	0.154	0.163	0.157	0.176	**0.147***	**0.155***	**0.152***	**0.155***
ITEA/DA	0.143	0.150	0.152	0.154	0.146	0.164	0.171	0.197	0.146	0.160	0.157	0.169	0.141	0.141	0.144	0.145
TRV/OI	0.153	0.156	0.152	0.160	0.158	0.166	0.169	0.194	0.156	0.165	0.162	0.187	0.146	0.150	0.151	0.154
TRV/EF	0.158	0.159	0.152	0.162	0.159	**0.174***	0.165	0.190	0.161	0.165	0.162	0.183	0.149	0.154	0.151	0.152
TRV/AC	0.158	0.160	0.157	**0.169***	0.163	0.168	0.170	0.195	0.157	0.155	0.155	0.167	**0.162***	0.157	**0.156***	0.160
TRV/OL	0.157	0.160	0.154	**0.171***	0.163	0.168	0.170	0.196	0.160	0.167	0.160	0.182	0.160	0.157	**0.156***	0.159
TRV/Avg	0.158	0.160	0.154	**0.170***	0.163	0.168	0.170	0.196	0.160	0.167	0.160	0.182	**0.160***	0.157	**0.156***	0.160
TRV/DA	**0.162***	0.160	**0.157***	**0.169***	0.163	**0.173***	0.169	0.186	**0.167***	0.162	0.167	0.186	**0.153***	0.157	0.157	0.160
TRAVOS+	0.159	0.161	0.155	0.165	0.161	0.172	0.173	0.186	**0.166***	0.162	0.168	0.185	**0.153***	0.157	0.157	0.159
ACT/OI	0.153	0.174	0.189	0.187	0.161	0.181	0.198	0.251	0.161	0.177	0.184	0.206	0.141	0.145	0.145	0.156
ACT/EF	0.147	**0.149**	**0.147**	**0.153**	**0.149**	0.157	**0.166**	**0.168**	0.149	**0.153**	**0.160**	**0.162**	0.145	0.146	0.145	0.148
ACT/AC	0.152	0.164	**0.164**	0.172	0.156	0.177	**0.176**	0.202	0.160	0.179	**0.167**	0.179	**0.156***	**0.154***	0.153	0.160
ACT/OL	0.151	**0.158**	**0.159**	0.163	0.156	0.173	**0.173**	0.182	0.159	**0.163**	**0.164**	0.166	**0.152***	0.153	0.154	0.159
ACT/Avg	0.151	**0.155**	**0.159**	0.162	0.154	0.173	**0.174**	0.181	0.158	**0.164**	**0.161**	0.162	**0.152***	0.153	0.153	0.157
ACT/DA	**0.148**	0.171	**0.176**	0.182	**0.154**	0.172	**0.188**	0.225	0.158	0.175	**0.177**	0.203	0.139	0.141	0.145	0.151
MET/OI	0.156	0.164	0.179	0.205	0.160	0.211	0.266	0.338	0.158	0.188	0.198	0.206	0.148	0.151	0.156	0.176
MET/EF	**0.149**	**0.148**	**0.155**	**0.162**	**0.149**	0.154	**0.171**	**0.186**	**0.150**	**0.156**	**0.169**	**0.172**	0.150	0.146	0.151	**0.151**
MET/AC	0.157	0.155	0.169	**0.172**	0.162	**0.169**	**0.183**	0.198	0.163	**0.170**	**0.177**	**0.187**	0.156	0.154	0.158	**0.155**
MET/OL	0.155	0.155	**0.169**	0.167	0.162	**0.165**	**0.181**	0.193	0.161	**0.161**	**0.172**	**0.174**	0.151	0.152	0.155	**0.155**
MET/Avg	0.155	0.155	**0.170**	**0.168**	0.161	**0.166**	**0.183**	0.196	0.161	**0.161**	**0.172**	**0.175**	0.152	0.151	0.157	**0.155**
MET/DA	0.151	0.171	0.180	0.203	0.163	0.202	0.253	0.318	0.163	0.186	0.199	0.201	0.147	0.150	0.158	0.171
	(a) Partly Random				(b) Fully Random				(c) Badmouthing				(d) Ballot-stuffing			

Table 1: Settings 1–4. OI = Only Indirect. EF = ExpFun. AC = ACT-RL. OL = Online Learning. Avg = Average. DA = Direct as Advisor.

	20%	40%	60%	80%	20%	40%	60%	80%	20%	40%	60%	80%	20%	40%	60%	80%
ITEA/OI	0.141	0.142	0.142	0.154	0.139	0.141	0.144	0.155	0.143	0.148	0.214	0.551	0.140	0.140	0.141	0.149
ITEA/EF	0.142	0.145	0.142	0.157	**0.149***	**0.147***	0.146	0.158	0.146	0.148	**0.192**	**0.506**	0.143	**0.144***	**0.147***	0.155
ITEA/AC	0.146	**0.151***	**0.158***	**0.171***	**0.148***	**0.149***	0.149	0.155	0.146	**0.158***	0.179	0.264	**0.146***	**0.148***	**0.148***	0.151
ITEA/OL	0.146	**0.154***	**0.156***	**0.165***	**0.146***	**0.147***	0.150	0.156	0.149	**0.159***	0.174	0.304	**0.147***	**0.146***	**0.148***	0.151
ITEA/Avg	0.146	**0.154***	**0.156***	**0.165***	**0.146***	**0.148***	0.148	0.156	0.148	**0.158***	0.174	0.299	**0.147***	**0.146***	0.146	0.151
ITEA/DA	0.140	0.143	0.143	0.158	0.139	0.141	**0.142**	**0.151**	0.143	0.146	0.201	**0.521**	0.140	0.142	0.141	**0.146**
TRV/OI	0.147	0.154	0.157	0.170	0.148	0.146	0.148	0.166	0.151	0.159	0.180	0.338	0.144	0.143	0.145	0.148
TRV/EF	0.152	0.156	0.154	0.169	0.148	0.150	0.151	0.157	0.151	0.160	0.171	0.322	0.147	0.145	0.146	0.152
TRV/AC	0.154	0.157	0.159	0.171	0.153	0.152	0.156	0.165	0.156	0.161	0.176	**0.252**	**0.159***	0.150	0.150	0.147
TRV/OL	0.154	0.157	0.159	0.171	0.153	0.151	0.154	0.160	0.156	0.161	0.175	**0.245**	**0.152***	0.148	0.151	0.149
TRV/Avg	0.154	0.157	0.159	0.171	0.153	0.151	0.154	0.160	0.156	0.161	0.175	**0.246**	0.152	0.146	0.149	0.148
TRV/DA	0.152	0.157	0.157	0.170	0.147	0.149	0.150	0.156	0.157	0.162	0.175	**0.309**	0.149	0.145	0.151	0.155
TRAVOS+	0.154	0.157	0.157	0.170	0.146	0.148	0.149	0.156	0.156	0.161	0.175	**0.310**	0.149	0.144	0.148	0.152
ACT/OI	0.174	0.222	0.257	0.319	0.140	0.143	0.147	0.161	0.176	0.262	0.481	0.631	0.140	0.141	0.147	0.186
ACT/EF	**0.142**	**0.141**	**0.146**	**0.157**	**0.146***	0.147	0.147	0.156	**0.143**	0.152	0.186	**0.512**	0.143	0.143	0.145	**0.153**
ACT/AC	**0.167**	**0.170**	**0.178**	**0.181**	**0.151***	**0.152***	**0.159***	0.172	**0.164**	0.188	0.207	0.220	**0.150***	**0.150***	0.156	**0.159**
ACT/OL	**0.147**	**0.154**	**0.160**	**0.171**	**0.148***	**0.151***	0.154	0.164	**0.151**	0.159	0.169	0.231	**0.148***	**0.148***	0.154	**0.153**
ACT/Avg	**0.147**	**0.150**	**0.155**	**0.165**	**0.148***	0.147	**0.156***	**0.176***	**0.151**	0.154	0.166	0.231	**0.149***	**0.149***	0.154	**0.153**
ACT/DA	0.176	**0.211**	**0.232**	**0.252**	0.141	0.143	0.144	0.158	0.175	**0.246**	**0.419**	**0.616**	0.141	0.142	0.146	**0.153**
MET/OI	0.151	0.155	0.166	0.210	0.150	0.155	0.166	0.211	0.162	0.228	0.350	0.507	0.153	0.168	0.244	0.387
MET/EF	0.149	0.151	0.165	0.203	0.149	0.155	0.166	**0.173**	**0.149**	0.162	0.219	0.294	0.148	**0.150**	**0.157**	**0.163**
MET/AC	**0.160***	**0.176***	0.176	**0.170**	0.158	0.163	0.177	**0.178**	0.162	**0.193**	0.210	0.230	0.158	0.156	0.171	0.171
MET/OL	0.157	**0.166***	0.175	0.198	0.155	0.162	0.162	**0.167**	0.158	**0.171**	0.195	0.243	0.155	0.157	**0.166**	0.165
MET/Avg	0.157	0.164	0.174	**0.182**	0.156	0.164	0.174	**0.184**	0.157	**0.169**	0.192	0.236	0.155	0.159	**0.169**	0.176
MET/DA	0.149	0.156	0.175	0.200	0.149	0.156	0.174	0.202	**0.155**	0.219	0.323	**0.450**	0.148	0.160	**0.216**	0.324
	(a) All-Negative				(b) All-Positive				(c) Selective BM				(d) Selective BS			

Table 2: Settings 7–10. OI = Only Indirect. EF = ExpFun. AC = ACT-RL. OL = Online Learning. Avg = Average. DA = Direct as Advisor.

	20%	40%	60%	80%	20%	40%	60%	80%	20%	40%	60%	80%	20%	40%	60%	80%
ITEA/OI	0.145	0.181	0.355	0.494	0.141	0.150	0.201	0.226	0.142	0.164	0.274	0.360	0.141	0.142	0.144	0.162
ITEA/EF	0.144	**0.144**	**0.144**	**0.144**	0.145	**0.144**	**0.144**	**0.144**	0.143	**0.140**	**0.142**	**0.148**	0.145	0.147	0.148	**0.151**
ITEA/AC	0.146	**0.147**	**0.163**	**0.191**	0.146	0.146	**0.151**	**0.158**	0.145	**0.146**	**0.148**	**0.157**	**0.157***	**0.158***	**0.161***	0.166
ITEA/OL	0.146	**0.146**	**0.146**	**0.149**	0.146	0.146	**0.146**	**0.149**	0.146	**0.146**	**0.147**	**0.153**	**0.152***	**0.157***	**0.157***	0.165
ITEA/Avg	0.145	**0.146**	**0.146**	**0.148**	0.145	0.146	**0.146**	**0.148**	0.145	**0.146**	**0.147**	**0.151**	**0.152***	**0.157***	**0.157***	0.161
ITEA/DA	0.142	**0.152**	**0.284**	**0.475**	0.140	**0.143**	**0.185**	**0.223**	**0.146***	0.161	**0.243**	0.350	0.140	0.142	**0.142**	**0.151**
TRV/OI	0.143	0.145	0.155	0.202	0.143	0.144	0.145	0.156	0.143	0.143	0.149	0.176	0.148	0.150	0.153	0.155
TRV/EF	0.146	0.146	0.146	**0.146**	0.146	0.146	0.146	**0.146**	0.146	0.147	0.146	**0.148**	0.156	0.156	0.157	**0.169***
TRV/AC	0.149	0.149	0.151	**0.151**	0.149	0.149	0.149	0.150	0.148	0.149	0.149	**0.154**	**0.163***	**0.164***	**0.176***	**0.186***
TRV/OL	0.149	0.149	0.149	**0.149**	0.149	0.149	0.149	0.149	**0.149***	**0.150***	0.149	**0.153**	**0.161***	**0.161***	**0.170***	**0.183***
TRV/Avg	0.149	0.149	0.149	0.149	0.149	0.149	0.149	0.149	0.148	0.149	0.148	**0.151**	**0.161***	**0.162***	**0.170***	**0.183***
TRV/DA	0.147	0.148	0.147	**0.147**	0.147	0.147	0.147	**0.147**	**0.149***	0.149	0.147	**0.151**	0.157	**0.161***	**0.166***	**0.180***
TRAVOS+	0.147	0.147	0.147	**0.147**	0.147	0.147	0.147	**0.147**	0.147	0.147	0.147	**0.150**	**0.160***	0.155	0.162	**0.173***
ACT/OI	0.156	0.222	0.426	0.581	0.141	0.149	0.215	0.238	0.148	0.187	0.329	0.407	0.161	0.173	0.198	0.271
ACT/EF	**0.143**	**0.145**	**0.146**	**0.143**	0.143	0.145	**0.146**	**0.143**	0.145	**0.142**	**0.149**	**0.149**	**0.143**	**0.147**	**0.144**	**0.156**
ACT/AC	0.151	**0.182**	**0.184**	**0.220**	**0.148***	0.147	**0.154**	**0.160**	0.151	**0.158**	**0.197**	**0.219**	0.150	0.152	0.163	0.189
ACT/OL	0.148	**0.147**	**0.150**	**0.150**	0.147	0.146	**0.149**	**0.149**	0.148	**0.146**	**0.150**	**0.156**	0.149	0.151	0.150	0.152
ACT/Avg	0.148	**0.147**	**0.151**	**0.153**	0.147	0.146	**0.148**	**0.149**	0.148	**0.147**	**0.150**	**0.158**	0.150	0.149	0.149	0.152
ACT/DA	0.151	**0.198**	**0.370**	**0.533**	0.142	**0.144**	**0.197**	**0.233**	0.147	**0.167**	**0.277**	**0.379**	**0.155**	**0.167**	**0.189**	**0.234**
MET/OI	0.183	0.368	0.588	0.773	0.156	0.189	0.223	0.248	0.169	0.271	0.385	0.472	0.159	0.193	0.248	0.336
MET/EF	**0.148**	**0.148**	**0.150**	**0.150**	**0.147**	0.147	**0.148**	**0.149**	**0.146**	**0.149**	**0.151**	**0.156**	**0.143**	**0.145**	**0.149**	**0.151**
MET/AC	**0.159**	**0.195**	**0.225**	**0.233**	0.157	**0.163**	**0.170**	**0.169**	**0.158**	**0.177**	**0.196**	**0.216**	0.152	**0.152**	**0.171**	**0.190**
MET/OL	**0.154**	**0.156**	**0.156**	**0.170**	**0.163***	0.163	0.163	0.163	**0.153**	**0.156**	**0.158**	**0.165**	0.151	0.146	0.152	0.153
MET/Avg	**0.153**	**0.158**	**0.166**	**0.201**	0.153	**0.155**	**0.158**	**0.170**	**0.154**	**0.158**	**0.164**	**0.188**	0.151	0.147	0.156	0.164
MET/DA	0.173	**0.327**	**0.544**	**0.716**	0.151	**0.179**	**0.215**	**0.238**	0.162	**0.249**	**0.360**	**0.442**	0.159	0.189	0.241	**0.311**
	(a) Camouflage				(b) Good-Bad-Good				(c) Periodic Behavior				(d) Random Behavior			

Table 3: Settings 11–14. OI = Only Indirect. EF = ExpFun. AC = ACT-RL. OL = Online Learning. Avg = Average. DA = Direct as Advisor.

	ExpFun	ACT-RL	Online Learning	Average	Direct as Advisor
1. Partly Random	**7, 0**	3, 3	5, 4	4, 3	2, 3
2. Fully Random	**10, 1**	5, 1	5, 1	5, 1	4, 1
3. BM	**8, 0**	5, 0	6, 0	6, 0	1, 1
4. BS	**1, 1**	1, 8	1, 6	1, 7	0, 1
5. Additive BM	**5, 0**	5, 3	5, 1	5, 1	3, 0
6. Additive BS	**3, 5**	1, 8	1, 5	1, 6	**0, 1**
7. All-Negative	**4, 0**	5, 5	4, 4	5, 3	3, 0
8. All-Positive	1, 3	1, 5	1, 4	1, 5	**2, 0**
9. Selective BM	**10, 0**	10, 1	10, 1	10, 1	7, 0
10. Selective BS	4, 2	4, 6	3, 6	3, 4	**4, 0**
11. Camouflage	**12, 0**	11, 0	11, 0	10, 0	10, 0
12. Good-Bad-Good	**10, 0**	7, 1	7, 1	7, 0	**10, 0**
13. Periodic	**11, 0**	**11, 0**	11, 2	**11, 0**	9,2
14. Random	**9, 1**	7, 7	8, 7	8, 7	7, 3

Table 4: Dominance analysis.

less than or equal to the corresponding value obtained by B, for each $q \in \{20, 40, 60, 80\}$, with at least one of the four inequalities being strict and statistically significant. Then, in all but three settings, no tested system is ever preferable to ITEA/ExpFun. Even the worst RFU values of ITEA/ExpFun exceed 0.165 in just three cases: 60% and 80% Selective BM, and 80% Fully Random Advisors.

Finally, we compared all four indirect trust systems, combined with ExpFun, to one another. ITEA/ExpFun had a total of 38 significant wins and only 5 significant losses, across all our experiments. ACT/ExpFun performed nearly as well, with 30 wins and 3 losses. By comparison, TRV/ExpFun had 9 wins and 35 losses, and MET/ExpFun had 7 wins and 41 losses. This suggests some special interplay between ExpFun and ACT, since ACT's rather inferior performance in terms of indirect trust is so greatly boosted by ExpFun.

6 Conclusions

We provided the first systematic study on when and how to combine direct with indirect trust in decision-making. The results of our broad empirical analysis show that the best methods for computing indirect trust benefit from incorporating direct trust only in certain categories of settings, especially when advisors change their behavior dynamically.

One of the methods for combining direct and indirect trust dominates all other tested methods, regardless of the indirect trust method used in conjunction. Combining this method with the indirect trust system ITEA yields a system that is very robust across a wide variety of scenarios and in most cases outperforms or is on par with all other tested systems.

References

[Cohen *et al.*, 2018] Robin Cohen, Peng F. Wang, and Zehong Hu. Revisiting public reputation calculation in a personalized trust model. In *Proceedings of the 20th International Trust Workshop*, pages 13–24, 2018.

[Huynh *et al.*, 2006] Trung Dong Huynh, Nicholas R. Jennings, and Nigel R. Shadbolt. An integrated trust and reputation model for open multi-agent systems. *Autonomous Agents and Multi-Agent Systems*, 13(2):119–154, 2006.

[Irissappane and Zhang, 2017] Athirai A. Irissappane and Jie Zhang. Filtering unfair ratings from dishonest advisors in multi-criteria e-markets: a biclustering-based approach. *Autonomous Agents and Multi-Agent Systems*, 31:36–65, 2017.

[Jiang *et al.*, 2013] Siwei Jiang, Jie Zhang, and Yew-Soon Ong. An evolutionary model for constructing robust trust networks. In *Proceedings of the 12th International Conference on Autonomous Agents and Multi-Agent Systems*, pages 813–820, 2013.

[Jøsang and Ismail, 2002] Audun Jøsang and Roslan Ismail. The Beta Reputation System. In *Proceedings of the 15th Bled Electronic Commerce Conference*, pages 2502–2511, 2002.

[Jøsang *et al.*, 2007] Audun Jøsang, Roslan Ismail, and Colin Boyd. A survey of trust and reputation systems for online service provision. *Decision Support Systems*, 43:618–644, 2007.

[Liu *et al.*, 2011] Siyuan Liu, Jie Zhang, Chunyan Miao, Yin-Leng Theng, and Alex C. Kot. iCLUB: an integrated clustering-based approach to improve the robustness of reputation systems. In *Proceedings of the 10th International Conference on Autonomous Agents and Multi-Agent Systems*, pages 1151–1152, 2011.

[Liu *et al.*, 2017] Yuan Liu, Jie Zhang, Quanyan Zhu, and Xingwei Wang. CONGRESS: A hybrid reputation system for coping with rating subjectivity. *IEEE Transactions on Computational Social Systems*, pages 163–178, 2017.

[Parhizkar *et al.*, 2019] Elham Parhizkar, Mohammad Hossein Nikravan, and Sandra Zilles. Indirect trust is simple to establish. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence*, pages 3216–3222, 2019.

[Regan *et al.*, 2006] Kevin Regan, Pascal Poupart, and Robin Cohen. Bayesian reputation modeling in e-marketplaces sensitive to subjectivity, deception and change. In *Proceedings of the AAAI National Conference on Artificial Intelligence*, pages 1206–1212, 2006.

[Teacy *et al.*, 2006] W. T. Luke Teacy, Jigar Patel, Nicholas R. Jennings, and Michael Luck. TRAVOS: Trust and reputation in the context of inaccurate information sources. *Autonomous Agents and Multi-Agent Systems*, 12:183–198, 2006.

[Teacy *et al.*, 2012] W. T. Luke Teacy, Michael Luck, Alex Rogers, and Nicholas R. Jennings. An efficient and versatile approach to trust and reputation using hierarchical Bayesian modelling. *Artificial Intelligence*, 193:149–185, 2012.

[Weng *et al.*, 2010] Jianshu Weng, Zhiqi Shen, Chunyan Miao, Angela Goh, and Cyril Leung. Credibility: How agents can handle unfair third-party testimonies in computational trust models. *IEEE Transactions on Knowledge and Data Engineering*, 22:1286–1298, 2010.

[Yu and Singh, 2003] Bin Yu and Munindar P. Singh. Detecting deception in reputation management. In *Proceedings of the 2nd International Conference on Autonomous Agents and Multi-Agent Systems*, pages 73–80, 2003.

[Yu *et al.*, 2014] Han Yu, Zhiqi Shen, Chunyan Miao, Bo An, and Cyril Leung. Filtering trust opinions through reinforcement learning. *Decision Support Systems*, 66:102–113, 2014.