Toward a Probabilistic Model of Trust in Agent Societies

Federico Bergenti

Dipartimento di Matematica
Università degli Studi di Parma
Viale G. P. Usberti, 53/A – 43100 Parma, Italy
federico.bergenti@unipr.it

Abstract. The literature about trust in societies of agents collects a huge number of works that analyse almost any facets of this concept from nearly every point of view. Nevertheless, an accepted and stable formal model of trust in agent societies is lacking. In this paper, we address this remarkable flaw of the current research by introducing a probabilistic model of trust capable of capturing two-party interactions, either direct of mediated by a Guarantor. Some interesting properties of this model are demonstrated and the final result of this work is an estimation (upper-bound) of the improvements that we expect from the inclusion of a Guarantor in a two-party interaction. In details, after an introductory section, Section 2 provides the foundations of our model and quantifies the increment of the utility that agents perceive because of the mediation of a Guarantor. Then, Section 3 deals with the decision-making strategies of rational agents and it shows a worst-case specialization of our model that justifies why agents are more likely choosing Guarantor-mediated interactions. Section 4 describes the overall results of this work in terms of bounds and evaluation of performances of the effects of mediation in interactions. Finally, Section 5 summarizes the main outcome of this work and outline some future lines of development.

1 Introduction

Interaction is a key feature of agenthood (maybe "the" key feature) and secure, trusted and privacy-aware interactions are what we truly want from real-world agent societies [4]. While it is easy to identify a minimum set of requirements capable of providing guarantees of security in multi-party interactions, e.g., authorization and authentication, we are not yet ready to identify similar requirements for trusted and privacy-aware interactions.

This work is along the lines of the research that is trying to identify a set of abstractions and mechanisms to guarantee trust and privacy-awareness in multi-agent interactions. In particular, the objective of this work is to develop a quantitative and probabilistic model of trust in order to show a sound proof of the convenience of Guarantor-mediated interactions over direct interactions. This objective is addressed taking into account a toy scenario that counts agent

X and agent Y only (two-party interaction). X is interested in signing a contract with Y and it is in the process of deciding whether to do it directly or through the mediation of a middleman, the Guarantor G. We take a rational standpoint and we assume that X discriminates between direct and mediated interaction on the basis of a utility function. Moreover, we take an incomplete information assumption and we say that X cannot take a fully-informed decision; rather it has to deal with a risky situation, which immediately turns our model into a probabilistic one.

The main results of the study of this toy scenario are based on a worst-case analysis of a much more general model and they provide a sound proof of why rational agents are more likely choosing Guarantor-mediated interactions rather then direct interactions.

This paper is organized as follows: next section provides the foundations of our model and it quantifies the increment of the utility that agents perceive after the inclusion of a Guarantor in a two-party interaction. Section 3 deals with the decision-making strategies of rational agents and it quantitatively shows why agents are more likely choosing Guarantor-mediated interactions. Section 4 describes the result of this work in terms of bounds and evaluation of performances of the effects of mediation in interactions. Finally, Section 5 summarizes the overall results of this work and outline some future lines of development.

2 A Model of Guarantor-Mediated Interactions

This section presents a set of abstractions and accounts for their relationships in order to setup a probabilistic model of interactions between agent X, agent Y and (possibly) Guarantor G. It is worth noting that this model is symmetrical for X and Y.

2.1 Abstractions

A very basic assumption that we take in the discussion of our model is that, from the point of view of security, trust and privacy, we can always reduce a two-party interaction to the act of signing of a contract. Therefore, from now on, we always refer to the joint act of signing a contract as a means to study any other form of two-party interaction.

Trust The problem of providing a quantitative definition of trust in societies of rational agents has been addressed in many different ways [14]. While we recognize the importance of cognitive models of trust, e.g., [5], we date back to the abstract and coarse-grained definition of trust given in [9] to come to a probabilistic interpretation this notion.

In particular, if *"Trust is the subjective probability by which an individual, A, expects that another individual, B, performs a given action on which its welfare depends,"* it is quite reasonable to model trust as an estimation of the probability by which B will perform the target action. Many factors contribute to this estimation [11, 13]; nonetheless we prefer to discard all these factors and we adopt a blackbox approach in which we model trust as a random variable $\mathbf{t}$ in an interval $[t_{min}, t_{max}]$.

The only assumption that we take in our model is that we require such an estimation to be performed by a rational agent A with some reasonable amount of information regarding B and its intentions of performing the action. This guarantees that the real probability of B performing the action lays in $[t_{min}, t_{max}]$, with both t_{min} and t_{max} reasonably strict around it.

Our model of two-party interactions relies on the following quantities, where X and Y are agents and c is a contract:

- $p_{c,X}$: the probability that X would carry out successfully all the obligations stated in c.
- $t_{c,X,Y}$: the level of trust X has in Y with respect to c, i.e., an estimation of $p_{c,Y}$ from the point of view of X.

Since trust expresses the estimation of a probability, it is clear that t_{min} and t_{max} are both between zero and one. The assumption $t_{max} \geq t_{min}$ is not restrictive.

Contract The study of all different forms of contract is subject of a large literature and even restricting to the types of contract that we normally consider in societies of agents [3], the diversity of possibilities is impressive. We acknowledge this literature, but for the sake of simplicity and for the need of quantitative tractability, we stick on a very simple model of contract. This model involves only two signers, X and Y, and it is totally described by two triples: each signer knows only one of the two triples.

From the point of view of agent X (but the notation is symmetrical for Y), a contract c is described by a triple, that we call *subjective evaluation*, that contains:

- A *reward* $R_{c,X}$ that agent X receives upon success of contract c;
- An *investment* $I_{c,X}$ that agent X makes in contract c, i.e., a certain assured value that it gives up when signing contract c; and
- A *penalty* $P_{c,X}$ that agent X receives if the contract fails because of the other party.

Such values are not restricted to be monetary quantifications, rather they quantify of the level of satisfaction of X. All in all, such quantities are subjective and therefore we cannot assess any mathematical relations between values of the triples of two different agents, even though they refer to the same contract.

More in details, a contract c has the following properties from the point of view of X:

- If the contract is honoured, agent X will receive $R_{c,X}$ with probability one; and
- If the contract fails because of agent Y, agent X will receive $P_{c,X}$ with probability one.

Another assumption concerns the relative ordering of reward, investment and penalty in a single subjective evaluation. We are interested in contracts whose parameters are ordered as follows:

$$P_{c,X} \leq I_{c,X} \leq R_{c,X}$$

This inequality captures the essence of risky contracts. Moreover, it implies that we are interested in agents that sign contracts with the intent of honouring them. Any failure in honouring a contract turns into a loss of utility (see later on): $I_{c,X} - P_{c,X}$. Furthermore, agents in our model do not consider their failure in honouring a contract, they assume that they can honour all contracts they sign; nevertheless the uncertainty about the other signer remains.

Guarantor The abstraction of Guarantors was introduced and discussed in details in [2, 1]. For the sake of completeness, we can simply say that here Guarantors are sources of highly trusted information and they are trust catalysts. If agent X requests a piece of information to Guarantor G, it assigns a correctness probability of one to the received response. Nevertheless, we introduce some failure probability in order to account for the idea that the use of additional information, i.e., the information that Guarantor provides, always introduces some risk, even though the information source is highly trusted and reliable.

2.2 Expectation of the Utility of Agents

The rest of this section analyses the utility that agents estimate in the process of signing a contract. This utility is formalised with and without the mediation of a Guarantor, and then such two cases are compared.

Direct interaction. Taking into account the previous definitions, we can explicitly write the expected value of the utility that agent X receives from a contract with agent Y:

$$\overline{U}^r_{X,c} = R_{c,X} \cdot p_{c,Y} + P_{c,X} \cdot (1 - p_{c,Y})$$

where the superscript "r" indicates that the real probability is used in this equation, and not an estimation of its value. This utility is not available to any agent since $p_{c,Y}$ is not observable. Instead, agent X estimates the expected utility using its trust in the other party (agent Y):

$$\overline{U}^e_{X,c} = R_{c,X} \cdot t_{c,X,Y} + P_{c,X} \cdot (1 - t_{c,X,Y})$$

Taking into account that agent X invests a certain value when it signs the contract, and that any contract has some probability $p^s_{c,X}$ of being finally signed, the total average utility that agent X perceives is:

$$\begin{aligned}\overline{U}^r_X &= \overline{U}^r_{X,c} \cdot p^s_{c,X} + I_{c,X} \cdot (1 - p^s_{c,X}) = \\ &= [R_{c,X} \cdot p_{c,Y} + P_{c,X} \cdot (1 - p_{c,Y})] \cdot p^s_{c,X} + I_{c,X} \cdot (1 - p^s_{c,X})\end{aligned}$$

As before, the agent can only estimate the total utility, obtaining:

$$\begin{aligned}\overline{U}^e_X &= \overline{U}^e_{X,c} \cdot p^s_{c,X} + I_{c,X} \cdot (1 - p^s_{c,X}) = \\ &= [R_{c,X} \cdot t_{c,X,Y} + P_{c,X} \cdot (1 - t_{c,X,Y})] \cdot p^s_{c,X} + I_{c,X} \cdot (1 - p^s_{c,X})\end{aligned}$$

Guarantor-mediated interaction. We can adapt the previous equations to the case in which the contract is evaluated using additional information obtained from a Guarantor.

In this case, the failure probability that we associate with a Guarantor has to be considered. This failure probability accounts for the possible uncertainty of the information that the Guarantor provides.

In particular, we assume that an error of a Guarantor may cause a failure of the contract. In this case agent X receives $P_{c,X}$. This risk is acceptable if we assume that in the case of an error, the Guarantor itself, and not contractors, pays the penalty.

Under this assumption, the new expected value of the utility of signing contract c is:

$$\overline{U}^{G,r}_{X,c} = R_{c,X} \cdot P\{c \text{ honoured}\} + P_{c,X} \cdot P\{c \text{ not honoured}\}$$

where the G superscript indicates that some information from the Guarantor is considered when signing the contract.

Under the assumption that p^G_k is the probability of the Guarantor to provide erroneous information and that any error of the Guarantor immediately causes the contract to fail, it is possible to express the total contract success and failure probabilities:

$$P\{c \text{ honoured}\} = p_{c,Y} \cdot p^G_k \tag{1}$$

$$\begin{aligned}P\{c \text{ not honoured}\} &= P\{c \text{ not honoured}|\text{Guarantor succeeds}\} + \\ &+ P\{c \text{ not honoured}|\text{Guarantor fails}\} \\ &= (1 - p_{c,Y}) \cdot p^G_k + 1 - p^G_k = 1 - p_{c,Y} \cdot p^G_k \end{aligned} \tag{2}$$

This allows rewriting the previous Equation (1) as:

$$\overline{U}^{G,r}_{X,c} = R_{c,X} \cdot p_{c,Y} \cdot p^G_k + P_{c,X} \cdot (1 - p_{c,Y} \cdot p^G_k)$$

Then, exploiting this equality in (1) we obtain the total average utility of signing the contract using information from a Guarantor as:

$$\begin{aligned}\overline{U}^{G,r}_X &= \overline{U}^{G,r}_{X,c} \cdot p^s_{c,X} + I_{c,X} \cdot (1 - p^s_{c,X}) = \\ &= [R_{c,X} \cdot p_{c,Y} \cdot p^G_k + P_{c,X} \cdot (1 - p_{c,Y} \cdot p^G_k)] \cdot p^s_{c,X} + \\ &\quad + I_{c,X} \cdot (1 - p^s_{c,X})\end{aligned}$$

Since agents give a trust of one to their Guarantors, most of the estimations of agent X are not changed by the mediation. In particular, the estimation of the contract success probability remains unchanged; therefore the estimation of the average utility of the contract does not change. Also, the estimation of the expected utility as a function of the probability of signing (1) is not influenced. As explained later on, the mediation of the Guarantor influences only the decision making strategy.

3 Decision Making Strategy

In this section, we introduce a rationality principle in our model by means of a decision making strategy that exploits utility to discriminate on the inclusion of the mediation of a Guarantor into an interaction.

3.1 Probability Density Function of Trust and the Risk Factor

As we said in Section 2, we model trust from the point of view of an agent as the estimation of the probability of having a contract honoured by its counterpart. An underlying assumption of this definition is that this estimation, and the real probability of the contract being honoured, both lie in the interval $[t_{min}, t_{max}]$. In essence, trust is a random variable $\mathbf{t}$ whose Probability Density Function (PDF) depends on decision making strategies of the agents involved in the contract.

Taking the variable $\mathbf{t}$ and a rationality principle into account, it is easy to define the probability that agent X would sign a given contract c.

In particular, this reasonable rationality principle mentioned above states that:

X decides to sign a contract c with Y if the estimated expected utility that it perceives is greater than the investment required to sign the contract

Which, in formal terms, is:

$$\begin{aligned} p^s_{c,X} &\doteq P\{\overline{U}^e_{X,c} > I_{c,X}\} = \\ &= P\{R_{c,X} \cdot t_{c,X,Y} + P_{c,X} \cdot (1 - t_{c,X,Y}) > I_{c,X}\} \end{aligned}$$

A further simple elaboration of this equation yields:

$$\begin{aligned} p^s_{c,X} &= P\{t_{c,X,Y} \cdot (R_{c,X} - P_{c,X}) > I_{c,X} - P_{c,X}\} = \\ &= P\left\{t_{c,X,Y} > \frac{I_{c,X} - P_{c,X}}{R_{c,X} - P_{c,X}}\right\} \end{aligned}$$

Where we supposed that $R_{c,X} - P_{c,X}$ is not zero. Now, if we define:

$$\kappa_{c,X} \doteq \frac{I_{c,X} - P_{c,X}}{R_{c,X} - P_{c,X}}$$

it is possible to express $p^s_{c,X}$ as:

$$p^s_{c,X} = P\{t_{c,X,Y} > \kappa_{c,X}\} \tag{3}$$

This last equation indicates that agent X signs contract c if its trust in the counterpart with respect to c exceeds $\kappa_{c,X}$, that we call *risk factor*. This factor depends only on X's subjective evaluation of contract c and it describes the risk that X perceives in signing contract c. This, allows to rephrase the decision making strategy as:

Agent X signs a contract c with a counterpart Y if and only if its trust in Y for contract c is greater then the risk factor of c.

It is worth noting that risk factor $\kappa_{c,X}$ is a number between zero and one. Furthermore, it is the quotient of two quantities that have a precise meaning on their own:

– The numerator $N_{c,X} = I_{c,X} - P_{c,X}$ expresses the gain that agent X obtains when rejecting contract c, in comparison to the case in which the contract is accepted but actually not honoured.
– The denominator $H_{c,X} = R_{c,X} - P_{c,X}$ represents the gain that the contract yields in case of success with respect to failure.

Then, e.g., if we consider the boundary cases:

– $\kappa_{c,X} = 1$ means that the contract will never be signed, because the investment equals the utility, but the first is guaranteed while the second is not.
– $\kappa_{c,X} = 0$ means that the contract has no risk, since the investment equals the penalty (which is assured with probability one). Therefore the contract will always be accepted.

In particular, if $\kappa_{c,X} \leq t_{min}$ the contract is always rejected, while if $t_{max} \leq \kappa_{c,X}$ the contract is always accepted. This consideration accounts also for the boundary cases analysis exposed above.

Having introduced the risk factor $\kappa_{c,X}$, it is possible to rewrite Equation (1) putting some emphasis on it. In particular:

$$\begin{aligned}\overline{U}^r_X &= [R_{c,X} \cdot p_{c,Y} + P_{c,X} \cdot (1 - p_{c,Y})] \cdot p^s_{c,X} + I_{c,X} \cdot (1 - p^s_{c,X}) = \\ &= [(R_{c,X} - P_{c,X})p_{c,Y} + P_{c,X}] \cdot p^s_{c,X} + I_{c,X}(1 - p^s_{c,X}).\end{aligned}$$

Now, explicitly showing $p^s_{c,X}$ and subsequently $(R_{c,X} - P_{c,X})$:

$$\begin{aligned}\overline{U}^r_X &= [(R_{c,X} - P_{c,X})p_{c,Y} + P_{c,X} - I_{c,X}] \cdot p^s_{c,X} + I_{c,X} = \\ &= (R_{c,X} - P_{c,X}) \cdot (p_{c,Y} - \kappa_{c,X}) \cdot p^s_{c,X} + I_{c,X}.\end{aligned}$$

This last equation gives the possibility to draw some interesting considerations. First, $\overline{U}^r_X$ is bounded between $P_{c,X}$ and $R_{c,X}$. Furthermore, $\overline{U}^r_X$ is a linear function of $p^s_{c,X}$, and its slope is $(R_{c,X} - P_{c,X}) \cdot (p_{c,Y} - \kappa_{c,X})$. Since $(R_{c,X} - P_{c,X})$

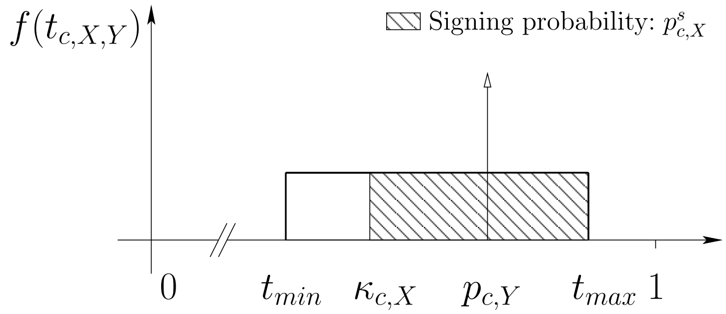

Fig. 1. The process of estimation of trust and its influence in decision making.

is non negative because of (2.1), the sign of the slope is influenced by $(p_{c,Y} - \kappa_{c,X})$ only. This ultimately means that the risk factor is an indicator of convenience in terms of average utility:

- If the success probability of the contract is greater than $\kappa_{c,X}$, then the average utility (of X) increases with the probability of signing the contract, i.e., the contract is advantageous.
- If the risk factor is lower than $\kappa_{c,X}$, the contract is disadvantageous and the average utility decreases with $p^s_{c,X}$.
- If $\kappa_{c,X} \equiv p_{c,Y}$, the average utility is constant.

3.2 Role of the PDF of Trust

The only working assumption that we took up to now is that $\mathbf{t}$ is a random variable bound by t_{min} and t_{max}. In this section, we further elaborate on trust as a random variable and, without breaking our blackbox approach, we go for the worst case and we assume that $\mathbf{t}$ is uniformly distributed in interval $[t_{min}, t_{max}]$. This new assumption allows us to study the influence of the mediation of a Guarantor on the average utility perceived by agents.

In accordance with Equation (3), we can express the signing probability as the probability that $t_{c,X,Y} \geq \kappa_{c,X}$. Therefore:

$$p^s_{c,X} = P\{t_{c,X,Y} > \kappa_{c,X}\} = \int_{\kappa_{c,X}}^{+\infty} f(t_{c,X,Y})\, \mathrm{d}t_{c,X,Y} \tag{4}$$

Then,

$$p^s_{c,X} = \begin{cases} 1 & \kappa_{c,X} \leq t_{min} \\ \frac{t_{max} - \kappa_{c,X}}{t_{max} - t_{min}} & t_{min} < \kappa_{c,X} < t_{max} \\ 0 & t_{max} \leq \kappa_{c,X} \end{cases} \tag{5}$$

Figure 1 shows a pictorial description of this last equation.

Now, we focus our analysis of the utility on the case in which $t_{min} \leq \kappa_{c,X} \leq t_{max}$, i.e., we exclude the edge cases. Moreover, we assume symmetric PDF of $\mathbf{t}$.

Introducing (5) in (4) we obtain the average utility as a function of t_{min} and t_{max}:

$$\overline{U}^r_X = (R_{c,X} - P_{c,X}) \cdot (p_{c,Y} - \kappa_{c,X}) \cdot \frac{t_{max} - \kappa_{c,X}}{t_{max} - t_{min}} + I_{c,X}. \qquad (6)$$

Then, using a symmetric PDF of $\mathbf{t}$ with width δ it is possible to rewrite (5) as:

$$p^s_{c,X} = \begin{cases} 1 & \kappa_{c,X} \leq t_{min} \\ \frac{t_{max}-\kappa_{c,X}}{t_{max}-t_{min}} & t_{min} < \kappa_{c,X} < t_{max} \\ 0 & t_{max} \leq k \end{cases} \qquad (7)$$

And then:

$$p^s_{c,X} = \begin{cases} 1 & \kappa_{c,X} \leq t_{min} \\ \frac{p_{c,Y}+\delta-\kappa_{c,X}}{2\delta} & t_{min} < \kappa_{c,X} < t_{max} \\ 0 & t_{max} \leq k \end{cases}$$

that expresses $p^s_{c,X}$ as a function of δ. Substituting this equation in Equation (6) and excluding the edge cases yields to:

$$\overline{U}^r_X = (R_{c,X} - P_{c,X}) \cdot (p_{c,Y} - \kappa_{c,X}) \cdot \frac{p_{c,Y} + \delta - \kappa_{c,X}}{2\delta} + I_{c,X}. \qquad (8)$$

This equation expresses the average utility as a function of the width of the probability density function of $\mathbf{t}$. Since the utility is a hyperbolic function of δ, any small decrease of δ implies a much higher increase in the average utility and vice versa. This introduces one of the main considerations of the paper:

If the information provided by a Guarantor linearly narrows the width of the PDF of $\mathbf{t}$*, then a hyperbolic increase of the average utility of having the contract signed occurs.*

Regarding the edge cases, when $\kappa_{c,X} \leq t_{min}$ or $t_{max} \leq \kappa_{c,X}$, the probabilities of signing the contract are one and zero respectively (5), and the utility is constant with respect to δ. Figure 2 shows the behaviour of the average utility for an increasing value of δ. The first plateau expresses the case in which $\kappa_{c,X} \leq t_{min}$, and consequently $p^s_{c,X}$ equals one.

Equation (8) has the following interesting consequence on the behaviour of the utility. If agent X takes its decisions on signing a contract c using a symmetric PDF for $\mathbf{t}$ centred in $p_{c,Y}$ and if the contract does not fail because of X, then $\forall \delta \in \mathbf{R} : \delta \geq 0,\ t_{min} - \delta \geq 0,\ t_{max} + \delta \leq 1, \overline{U}^r_X(\delta)$ is non-increasing. In fact, $\overline{U}^r_X$ is piecewise differentiable and the differentiation of (8) for $t_{min} < \kappa_{c,X} < t_{max}$ yields to:

$$\frac{\partial \overline{U}^r_X}{\partial \delta} = (R_{c,X} - P_{c,X}) \cdot (p_{c,Y} - \kappa_{c,X}) \cdot \frac{2\delta - 2(p_{c,Y} - \kappa_{c,X} + \delta)}{4\delta^2} =$$
$$= -\frac{(R_{c,X} - P_{c,X}) \cdot (p_{c,Y} - \kappa_{c,X})^2}{2\delta^2}$$

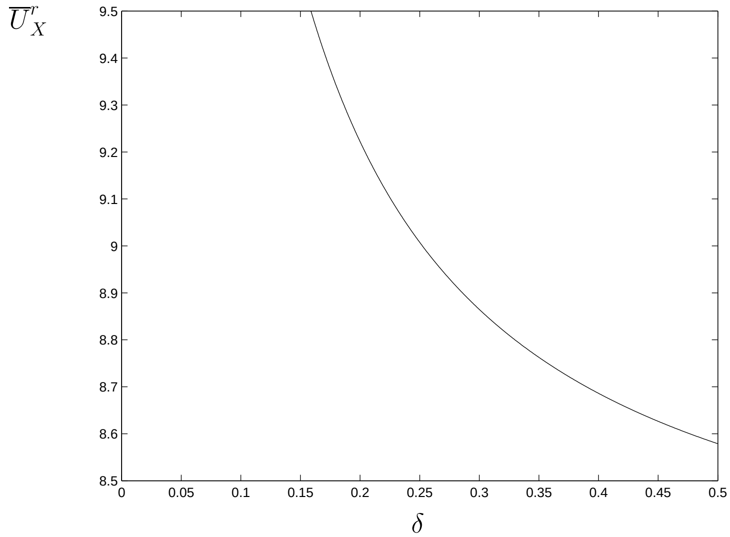

Fig. 2. Hyperbolic decrease of $\overline{U}^r_X(\delta)$.

Taking into account that a subjective evaluation is well formed if $R_{c,X} \geq P_{c,X}$, the partial derivative is always non-positive, i.e., any enlargement of the estimation (which introduces uncertainty), worsen the performance of the agent's decision strategy and its relative utility.

The explicit choice of a PDF for trust $\mathbf{t}$ allows elaborating on the inclusion of mediation into an interaction. The two parameters $\kappa_{c,X}$ and $p_{c,Y}$ are kept fixed, since the mediation of a Guarantor does not change or influence them. On the contrary, the total error probability is modified to account for the additional probability of error that the Guarantor brings. Using Equation (1), it is possible to directly substitute $p_{c,Y}$ with $p_{c,Y}\, p^G_k$ to express the total success and failure probabilities, thus obtaining the equivalent of (8) for the case of Guarantor-mediated interactions. To stress the fact that the width of the estimation is different when introducing a Guarantor in the interaction, we use δ^G instead of δ:

$$\overline{U}^{G,r}_X = \begin{cases} H_{c,X} \cdot M^G + I_{c,X} & \kappa_{c,X} \leq t_{min} \\ H_{c,X} \cdot M^G \cdot \frac{p_{c,Y} + \delta^G - \kappa_{c,X}}{2\delta^G} + I_{c,X} & t_{min} < \kappa_{c,X} < t_{max} \\ I_{c,X} & t_{max} \leq \kappa_{c,X} \end{cases} \tag{9}$$

Where we defined (see next section) $M^G = (p_{c,Y}\, p^G_k - \kappa_{c,X})$.

However, it is important to note that this estimation is always centred around $p_{c,Y}$, since agent X accords a trust of one to its Guarantor.

4 Results and Bounds

In this section we study the effects of mediation in our model. In order to do so, we recall that our working assumption is that Guarantors provide additional information to agents, thus allowing for a more precise (narrower) estimation of probability $p_{c,Y}$. Anyway, Guarantors, although highly reliable, introduce additional error probability, that must be compensated by improvements in the estimation of trust.

In order to quantify the performance of a Guarantor as a middleman in an interaction between agent X and Y, we calculate the amount of additional information that a Guarantor needs to provide in order to keep the average utility of agent X fixed.

The comparison of the two utilities expressed in Equations (8) and (9) allows to calculate the width of Guarantor-mediated estimation of trust for which the utility equals the case without mediation. If we introduce $\hat{\delta}^G$, a function of δ and p_k^G, as:

$$\hat{\delta}^G \doteq \delta_G \in \left[0, \frac{1}{2}\right] : \overline{U}^r_X(\,\delta\,,\,p_{c,Y}\,) = \overline{U}^{G,r}_X(\,\delta_G\,,\,p_{c,Y} \cdot p_k^G\,)$$

we can compare Equations (8) and (9) to obtain:

$$(p_{c,Y} - \kappa_{c,X}) \cdot \frac{p_{c,Y} + \delta - \kappa_{c,X}}{\delta} = (p_{c,Y}\,p_k^G - \kappa_{c,X}) \cdot \frac{p_{c,Y} + \hat{\delta}^G - \kappa_{c,X}}{\hat{\delta}^G}$$

where we subtracted $I_{c,X}$ on both sides and multiplied by $\frac{2}{R_{c,X} - P_{c,X}}$. Then, introducing $M = (p_{c,Y} - \kappa_{c,X})$ and $M^G = (p_{c,Y}\,p_k^G - \kappa_{c,X})$:

$$M \cdot \frac{\delta + M}{\delta} = M^G \cdot \frac{\hat{\delta}^G + M}{\hat{\delta}^G}$$

and dividing by M^G yields:

$$\frac{M}{M^G} \cdot \frac{\delta + M}{\delta} = \frac{\hat{\delta}^G + M}{\hat{\delta}^G}.$$

This last equation allows to make $\hat{\delta}^G$ explicit:

$$\hat{\delta}^G = \frac{M}{\frac{M}{M^G} \cdot \frac{\delta + M}{\delta} - 1} = \frac{M\,M^G\,\delta}{M(\delta + M)M^G\,\delta}$$

that holds if $M^G \neq 0$.

It should be quite clear that $\hat{\delta}^G$ is the breakeven point that makes agent X choose to go for a mediated interaction rather than for a direct interaction:

- If the Guarantor provides enough information to restrict the estimation of trust to a width less than $2\,\hat{\delta}^G$, the use of the mediation is advantageous.

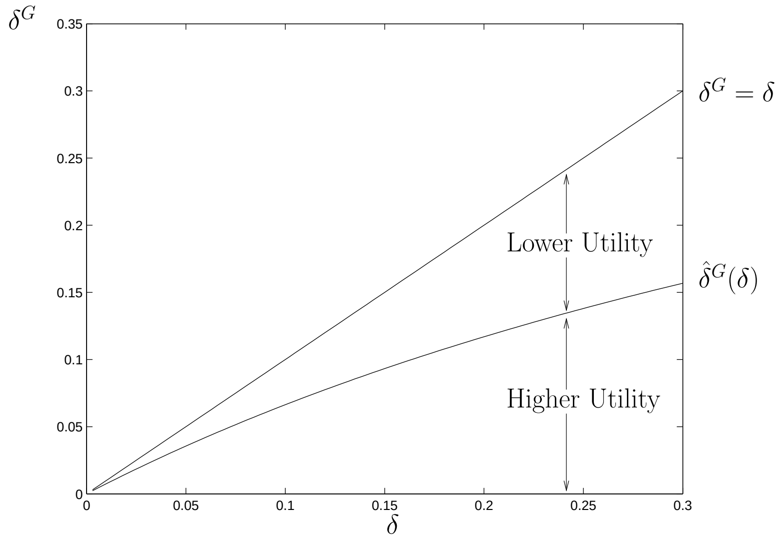

Fig. 3. Zone of convenience for mediated interactions.

- If the estimation remains larger than $2\,\hat{\delta}^G$, the error probability introduced by the Guarantor decreases the average utility.

It is worth noting that this decision strategy is purely ideal because agent X does not know p_k^G. Anyway, Figure 3 provides an ideal means for evaluating the zone of convenience for choosing mediated interactions.

In order to ground our model to everyday experience, we recall that we are interested in Guarantors that introduce a very low probability of error, and therefore we study the behaviour of $\hat{\delta}^G$ as p_k^G tends to one. What we obtain from this study is that if agent X makes its decisions using a symmetric PDF and that the contract does not fail because of X, $\forall \delta \in \mathbf{R} : 0 \leq \delta \leq \frac{1}{2}$, $\delta - \hat{\delta}^G$ tends to zero in an hyperbolic way as p_k^G tends to one. Due to the lack of space in this paper, we cannot go in the details of the demonstration of this result, anyway it is worth saying that the demonstration is divided into four steps:

- Proof that $\hat{\delta}^G$ is a hyperbolic function of p_k^G;
- Proof that $\hat{\delta}^G$ is increasing from a certain value of p_k^G on, excluding the edge cases;
- Proof that its maximum value is δ;
- Proof for the edge cases.

This result shows that if a Guarantor introduces a (sufficiently) low probability of error, the use of its mediation is advantageous and the advantages that it brings are fast increasing as the probability of error decreases. As a further evidence of the convenience of using Guarantor-mediated interaction, the behaviour of the zone of convenience with respect to p_k^G is shown in Figure 4.

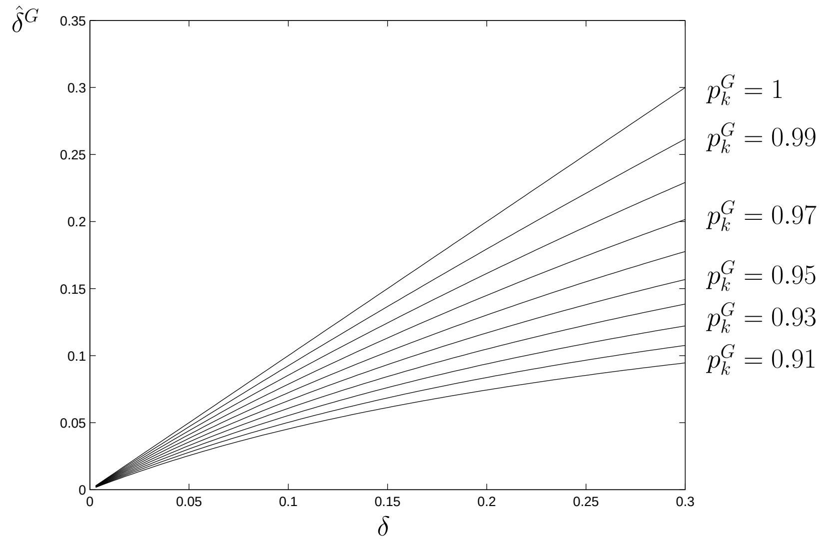

Fig. 4. $\hat{\delta}^G$ as a function of δ for different values of p_k^G.

5 Conclusions

The aim of this work is to provide a sound demonstration that the development of Guarantor-mediated infrastructures is extremely beneficial to support secure, trusted and privacy-aware interactions in real-world societies of agents. In particular, such infrastructures provide notable features that are not discussed here, but that play a fundamental role from the point of view of scalability, reliability and traceability (see [1, 2]). Then, in many cases, the additional utility that mediation provides to agents is considerable even through Guarantors are not error-free.

This work is not meant to be conclusive and many points remain open. One of the major planned developments regards the study of concrete trust estimators, and the introduction of the resulting PDFs in our model. Another very important open point regards the study of the effects of delegation of tasks and goals through a chain of delegated Guarantors.

Furthermore, the study of one of the main features of Guarantors, i.e., the possibility of anonymising interactions, is still in search of a formalization (and of a probabilistic model), even though its characteristics and possible uses are clearly understood [1]. This kind of interaction allows to prevent unwanted spread of sensible information; as such, its study remains central in the evaluation of the agent's benefits from the Guarantor infrastructure.

References

1. Bergenti, F., Bianchi, R. and Fontana, A. (2005). Secure and Trusted Interactions in Societies of Electronic Agents. In Proceedings of *The* 4^{th} *Workshop on the Law and Electronic Agents (LEA 2005)*, 1–12. Bologna, Italy. Wolf Legal Publishers.
2. Bianchi, R., Fontana, A and Bergenti, F. (2005). A Real–World Approach to Secure and Trusted Negotiation in MASs. In Proceedings of *The* 4^{th} *International Joint Conference on Agents and Multi-Agents Systems (AAMAS)*, 1163–1164. Utrecht. The Netherlands. ACM Press.
3. Bons, R.W.H. (1997). *Designing Trustworthy Trade Procedures for Open Electronic Commerce*. Ph.D.diss., EURIDIS and Faculty of Business Administration, Erasmus University, Rotterdam, The Netherlands.
4. CASCOM Web site. `http://www.ist-cascom.org`
5. Castelfranchi, C. and Falcone, R. (1998). Principles of Trust for MAS: Cognitive Anatomy, Social Importance, and Quantification. In Proceedings of *The International Conference of Multi-agent Systems (ICMAS)*, 72–79. Paris, France. ACM Press.
6. Debenham, J. and Sierra, C. (2005). An Information-based Model for Trust. Proceedings of *The* 5^{th} *International Conference on Autonomous Agents and Multiagent systems (AAMAS)*, 497-504, Utrecht, The Netherlands.
7. Ellison, C. (1999). *SPKI Requirements*. IETF RFC 2692.
8. Ellison, C., Frantz, B., Lampson, B., Rivest, R., Thomas, B., and Ylonen, T. (1999). *SPKI Certificate Theory*. IETF RFC 2693.
9. Gambetta, D. (Ed.) (1988).*Trust: Making and Breaking Co-operative Relations*. Basil Blackwell, Inc.
10. JENA Web site. `http://jena.sourceforge.net`
11. Jennings, N. R., Parsons, S., Sierra, C. and Faratin, P. (2000). Automated Negotiation. In Proceedings of *The* 5^{th} *International Conference on the Practical Application of Intelligent Agents and Multi-Agents Systems (PAAM-2000)*, 23–30. Manchester, UK.
12. Jøsang, A., Ismail, R. and Boyd, C. (2007). A Survey of Trust and Reputation Systems for Online Service Provision. *Decision Support Systems*, 43(2), 618–644
13. Marsh, S. (1994). *Formalising Trust as a Computational Concept*. Ph.D. diss., Department of Mathematics and Computer Science, University of Stirling, Stirling, UK.
14. MINDSWAP Team. *A Definition of Trust for Computing with Social Networks* Technical report, University of Maryland, College Park, February 2005.
15. OWL Web site. `http://www.w3.org/2004/OWL`
16. Racer Web site. `http://www.racer-systems.com`
17. Yu, B. and Singh, M. (2003). Searching Social Network. In Proceedings of *The* 2^{nd} *International Joint Conference On Autonomous Agents and Multiagent Systems*, ACM Press.

Acknowledgements

This work is partially supported by project CASCOM (FP6-2003-IST-2/511632). The CASCOM consortium is formed by DFKI (Germany), TeliaSonera AB (Sweden), EPFL (Switzerland), ADETTI (Portugal), URJC (Spain), EMA (Finland), UMIT (Austria), and FRAMeTech (Italy). The authors would like to thank all partners for their contributions.