

Copyright © 2007 IEEE. Reprinted from Second International Conference on Availability, Reliability and Security (ARES'07).

This material is posted here with permission of the IEEE. Such permission of the IEEE does not in any way imply IEEE endorsement of any of the University of Bergen's products or services. Internal or personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution must be obtained from the IEEE by writing to pubs-permissions@ieee.org.

By choosing to view this document, you agree to all provisions of the copyright laws protecting it.

Why Trust is not proportional to Risk

Bjørnar Solhaug^{1,2}, Dag Elgesem¹ and Ketil Stølen^{2,3}

¹Dep. of Information Science and Media Studies, University of Bergen

²SINTEF ICT, ³Dep. of Informatics, University of Oslo

Email: {bjors,kst}@sintef.no, dag.elgesem@infomedia.uib.no

Abstract

Trust is inherently related to risk, but for trust assessment to be integrated with the management of the risks involved in trust based cooperation, the exact relation must be well understood. Existing literature on trust management is neither clear nor unambiguous on this issue. This paper discusses notions of trust as presented within the disciplines of sociology and economics for the purpose of motivating trust management. A critical survey of state of the art literature on trust management is provided, where weaknesses and ambiguities with respect to clarifying the notion of trust are discussed. An analysis and explanation of the exact relationship between risk and trust is presented, and implications of the subjectivity of trust relations are accounted for.

1. Introduction

The term *trust management* was coined in 1996 by Matt Blaze who refers to it as a systematic approach to managing security policies, credentials and trust relationships for the purpose of making security critical decisions regarding authorization and delegation of access rights [5, 4]. Trust management has since then been the subject of increased attention, with the expense of today being a label for a diversity of approaches. In a more recent paper, trust management is described as an activity in the “intersection between sociology, commerce, law and computer science” [14].

There is nevertheless a shared ground to the various approaches, viz. the relation between *trust* on the one hand and *security* and *risk* on the other. Information security concerns the preservation of confidentiality, integrity and availability [13]. In short, this means, respectively, to ensure that unauthorized users cannot access information, that information is kept accurate and complete, and that information is kept accessible to authorized users. A risk can be understood as the probability of the occurrence of an (unwanted) incident [12] with a negative impact on an asset, i.e. some-

thing of value. The level of risk is given as a function from the consequence of the incident and the probability of its occurrence [2]. A security risk is the probability of an incident resulting from a security breach.

The diversity of approaches to trust management stems in part from differences in the context in which trust is managed and in part from differences in how trust management should be applied. This simply indicates that there are many different security domains in which trust plays a crucial role and should be dealt with systematically. The diversity that stems from different understandings of the very notion of trust, on the other hand, is of another character.

In a situation of trust there is always the possibility of deception or betrayal. The trusting party has certain expectations about the future behavior of the trusted party, however knowing that the trusted party has the freedom to disappoint the expectations. There is hence an inevitable relation between trust and risk; without risk it does not make sense to talk about trust [7, 19]. The exact relation between the two concepts is not clearly and unambiguously accounted for, though, and this has led to a more fundamental confusion about what trust management is all about. There is also the question about the extent to which trust at all contributes to the understanding and management of security needs. Risk management is quite well understood, and it is obvious that some of the existing approaches to managing trust can be understood in terms of risk management. In those cases trust is at best redundant, but may also contribute negatively by blurring the issues and adding to the confusion.

The objectives of this paper are on a conceptual level and on trust management foundations. We do not address trust problems for the IT domain per se, but we believe that the problems we discuss must be resolved in order to reach a proper understanding of trust within the IT domain.

We will in this paper first capture a notion of trust that can be transferred into the trust management domain by reviewing existing approaches within sociology and economics. In Section 3 we provide a critical survey of existing attempts to relate the notions of trust and risk, and point out weaknesses in these approaches. Section 4 presents

our own suggestions for clarifying the exact relationship between trust and risk. In Section 5 we explain the notion of well-founded trust and discuss issues involved in misplacing trust. Finally, in Section 6, we conclude.

2. Understanding Trust

Trust is understood very differently in the literature on the subject, which to a large extent is due to the many aspects of trust. A psychologist is likely to operate with an understanding that to some degree differs from the understanding of trust promoted by a sociologist. This is not to say that the one is wrong and the other right, it is more a question of identifying aspects of trust that are relevant for the particular domain. So is the case for trust management.

Trust management can be characterized as a special case of risk management with particular focus on authentication of and cooperation with actors the identity and/or intentions of which may be uncertain. The objective is, on the basis of trust assessments, to make decisions about actual or potential interactions with these actors.

This characterization of trust management is from the perspective of the trusting party. From the perspective of the trusted party, trust management is about increasing or maintaining, and correctly representing, its own trustworthiness [14]. In this paper, as in most of the literature, we will focus on the perspective of the trusting party.

The aspects of trust relevant to the trust management domain are foremost found within sociology and economics. Yet, approaches to trust management motivate the relevance of trust differently, an issue we will return to subsequently.

Our point of departure is works by Gambetta [8] and Williamson [24]. They both address the question of cooperation, which is highly relevant for trust management. Gambetta's definition of trust is much referred to in the trust management literature, see e.g. [1, 6, 14, 20], and his focus is on the motives actors may have for cooperation. Williamson analyzes many of the trust issues raised by Gambetta and shows that trust within impersonal relations, i.e. those not associated with relations of family, love, affection, friendship and the like, is best understood as the result of calculative reasoning. The results of Williamson are useful both as a clarification of aspects of the notion of trust and as a basis for trust management.

2.1. Trust as a Basis for Cooperation

Trust is a relationship between two entities, a *trustor* and a *trustee* where the former places trust in the latter. A particular relationship is valid in a specific set of contexts. I may for example trust my doctor to provide the correct diagnosis, but not to determine what is wrong with my broken car. Gambetta states that (emphasis original) "trust (or,

symmetrically, distrust) is a particular level of the subjective probability with which an agent assesses that another agent or group of agents will perform a particular action, both *before* he can monitor such an action (or independently of his capacity ever to be able to monitor it) *and* in a context in which it affects *his own* action" [8].

The crucial question for the trustor is whether or not to engage in a cooperation with the trustee, and this depends on the extent to which the trustor *believes* that the trustee will behave in a certain way. The level of trust is hence determined subjectively based on the evidence available to the trustor about the intentions of the trustee and the constraints regulating the trustee's choices of action.

There is of course an *objective* probability that the trustee will perform a particular action. This probability may not be known to the trustor, but it is in her interest to estimate this probability as accurately as possible. We shall say that *well-founded trust* is one in which the trustor *knows* the probability. As we will argue below, this knowledge is required for an exact risk estimation.

Following Gambetta, trust is a threshold point located on a probabilistic distribution between 0 and 1, where 0 corresponds to complete distrust and 1 to complete trust. When the probability is 0.5 the evidence available to the trustor gives no indication as to what is the intention of the trustee. Furthermore, "trust is particularly relevant in conditions of *ignorance* or uncertainty with respect to unknown or unknowable actions of other" [8] (emphasis original).

It seems here like Gambetta makes no distinction between ignorance and uncertainty. We believe it is a mistake not to make this distinction, and that it is a mistake of a kind that has contributed to obscure the relation between the notions of trust and risk. Our point is that there is an important difference between *knowing* that the probability for the occurrence of an event is 0.5 and being ignorant about it. In the case of well-founded trust, the trustor knows the risks and is in a position in which she can calculate the consequences of her choices of action. In the case of ignorance, she will either be in a situation in which there is lack of evidence and hence a lack of a basis upon which she can make decisions, or she will do her risk calculation and make decisions based on false information. In the latter case the risks will obviously be wrongly estimated.

We do follow Gambetta in that trust indeed is an issue in cases of ignorance; if we had a complete overview of all possible contingencies and their respective probabilities, trust would not be an issue. Where we differ from Gambetta is that he requires the level of trust to be located at the top or bottom end of the probability distribution for ignorance to be ruled out, whereas we stress the importance of well-foundedness.

The more constrained the actions of others, the easier it is to reach a well-founded estimate of their trustworthiness.

Gambetta puts this by stating that the domain of trust is one in which “agents have a degree of freedom to disappoint our expectations” [8]. Should heavy constraints remove this freedom, there would be no need for trust. We shall use the term *assurance* [6, 21] to refer to the various constraints decreasing the need for and relevance of trust. These may be laws or legal contracts, coercion, social norms, etc.

Gambetta analyzes trust in the context of decisions about whether or not to engage in a cooperation, and the level of trust is the determining factor. Although determining, before the a decision is made this factor must be related to the *consequence* of misplacing trust. Gambetta illustrates this with an example borrowed from Axelrod [3]: “to walk about in the trench in sight of the enemy requires an extremely high degree of trust that the enemy will observe the implicit truce, and the costs of being wrong may prove much more serious than those of lying low”. Hence, the more severe the consequence, the more demand on the value of the probability.

So far we have mostly addressed the question of deception in relation to trust. Equally important is the reason or motivation for someone to trust someone else and cooperate with him, viz. the opportunities that may be opened for. We will use the term *opportunity* as the dual to the notion of incident. Whereas the latter has a negative impact on an asset, the former has a positive one.

The value of a risk follows functionally from the consequence of the incident and the probability of its occurrence. The upper part of Fig. 1 shows a class diagram for the notion of risk. The class diagram for the dual to a risk, which we will refer to as a *prospect*, is given in the lower part of the figure. The prospect value is derived as a function from the positive consequence of an opportunity and the probability of the occurrence of the opportunity.

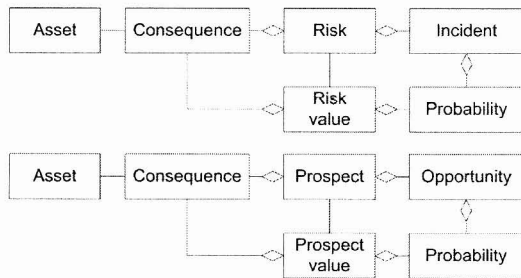


Figure 1. Risk and prospect

Gambetta observes that the cost of refraining from engaging in cooperation may be unacceptably high. This is partly due to the possible loss of opportunities. The decision making within the context of trust and cooperation is hence an exercise in calculating risks and prospects and acting on the basis of the derived results.

2.2. Trust as Calculative Reasoning

Williamson’s discussion of trust [24] is from the perspective of transaction cost economics (TCE). Two important assumptions on which TCE is based are *bounded rationality* and *opportunism*. The former refers to the limited capacity of people to memorize and calculate; there is no way neither to grasp all the possible contingencies at once, nor to determine the consequences of the information we do possess. Opportunism refers to what Williamson calls “self-interest seeking with guile” [24]. Not only do actors pursue their own interest, they may do this deceitfully and take advantage of circumstances that give them a chance to exploit others. The result of the two assumptions is to “organize transactions so as to economize on bounded rationality while simultaneously safeguarding them against the hazards of opportunism” [24].

With respect to trust, Williamson states that TCE “refers to contractual safeguards, or their absence, rather than trust, or its absence” [24]. In our own terms this is to say that the issue is to decide upon which transactions to do on the basis of the current assurance, not the basis of our trust in the other party. The crucial point is that if the relations of interest are of a calculative character, they should be described in calculative terms. In other words, trust is a diffuse notion of mixed meanings that should be avoided if possible.

It should be stressed that Williamson is concerned with *impersonal* relations of trust and that it is in this context that trust can be seen as a redundant notion that is better understood in terms of cost/benefit and calculated risk. The situation in which the affected parties can decide upon which transactions to make in a calculative manner is one in which they “(1) are aware of the range of possible outcomes and their associated probabilities, (2) take cost-effective actions to mitigate hazards and enhance benefits, (3) proceed with the transaction only if net gains can be projected, and, (4) if X can complete the transaction with any of several Ys, the transaction is assigned to that Y for which the largest net gain can be projected” [24].

Needless to say, this situation is quite an ideal one as the first requirement is not easily obtained. Bounded rationality is, however, one of the assumptions of the analysis, and the point is to understand decision making with respect to cooperation as a calculative exercise and in this way show that the notion of trust as captured by Gambetta [8] is redundant.

We believe that Williamson contributes to the clarification of the notion of trust, in particular within the domain of trust management where the relevant relations generally are those of an impersonal character and where there is a need to make calculated assessments and decisions in relation to risks. Williamson does not, however, address the problem stressed by Gambetta that trust is particularly relevant in conditions of ignorance. It is imperative to have a strategy

for how to deal with information shortage, and it is not obvious that the better solution is to do calculative reasoning on the basis of available information as if this information was complete.

There is nevertheless an agreement between the two in that risk is not all there is to trust. There is an aspect of prospect to it and the challenge is to not miss opportunities while simultaneously avoiding incidents to occur.

2.3. Managing Trust

The aim of the above subsections was to capture a notion of trust that can be transferred into and handled by the domains of risk and trust management. The works by Gambetta [8] and Williamson [24] motivate and justify the trust management activity of assessing trustworthiness, calculating risks and making rational choices on this basis.

An obvious discussion, which is outside the scope of this paper, is the extent to which the notions of trust as handled within trust management reflect the aspects of trust as is expressed more generally within the discipline of sociology. Some trust models emphasize that the very *absence* of conscious considerations is what characterizes trust and, furthermore, that there is no situation in which an individual can choose whether to trust or not. Trust does simply not fit into the rational choice framework, and is something that cannot be managed in the same way as for example risks are handled within risk management [11].

The remainder of this paper assumes a notion or an aspect of trust that fits into the framework of risk management. In the following sections we will elaborate on the precise relationship between trust and risk.

3. Existing Elucidations

It is not surprising that trust has become a subject within the domains of risk and security management. It is crucial to be able to correctly assess trust such that security critical decisions can be made on a well-founded basis. The exact relation between risk and trust is less obvious by judging from the literature on trust management.

Jøsang and Tran [16] emphasize the importance of making trust assessments of potential transaction partners for the purpose of minimizing risks. The basic idea seems to be that the more trustworthy a potential partner, the less risk involved in doing transactions with him. By defining a risk acceptance level, we should be able to determine whether or not to transact with a potential partner whom we trust to a specific degree, but neither a framework for such a decision making is provided, nor an explanation of the exact relation between trust and risk. It is in particular not obvious, as we will argue below, that high trustworthiness necessarily implies low risk.

Grandison and Sloman [9] address trust issues for internet applications, and also they observe that there is a relation between trust and risk, yet state that “there has been very little work on using risk management frameworks for trust management or on the analysis of the exact relationship between risk and trust”. What they do suggest, however, is that the “level of trust has an approximate inverse relationship to the degree of risk”. But it is certainly so that if one party trusts another party to a very high degree, she is willing to place high values at stake in transactions involving him. On the other hand, if she barely trusts him at all, she will place very low values at stake and thus avoiding risk. This inverse relation is hence not apparent.

In a later paper [10], Grandison and Sloman argue likewise by stating that “higher risk implies less trust”. Since they define risk as “a probability of failure” they may consistently say that there is such an inverse relation between trust and risk, but defining risk without including the damage of the failure gives a notion of risk that is not very useful. This definition is furthermore somewhat surprising as they do remark that “risk evaluation may depend on factors such as value of a transaction”.

Broadly speaking, risk management, see e.g. [2], seeks to identify risks and to determine whether those risks should be accepted, reduced or removed. Povey [22] understands trust management as seeking “to identify the circumstances under which we are prepared to accept risks that may be exposed by relying on certain entities”. By identifying risk as the common element of trust management and risk management, he claims that we may hold trust as a form of risk acceptance. The idea is that “you are prepared to accept risks if you trust the entities that can expose them” [22]. These claims witness that the understanding of how trust relates to risk is not clearly accounted for. Firstly, given a risk that is so severe that it will put a whole enterprise out of business, how can this risk be accepted even if the one introducing it is trusted? Secondly, does it make sense at all to say that someone introducing such extreme risks is trusted?

We will not go into details on diverse further approaches to this issue, only mention that there is a number of other influential contributions to the area of trust and trust management that are unclear about the precise relation between trust and risk, e.g. [18, 20, 23, 24]. This observation is not to question of the substance of these contributions, it is rather an indication of a need for a clarification. Before we provide our own suggestion, however, we need to mention the closely related work by Jøsang and Presti [15].

The purpose of their paper is very much the same as the purpose of this. They observe, as do Grandison and Sloman [9], that current trust systems and models generally fail to explicitly capture the relation between trust and risk. Also they seem to acknowledge the idea that there is an inverse relationship between the two notions. Despite this miscon-

ception, though, they do provide definite clarifications.

In their analysis they consider transactions with two possible outcomes, gain or loss. Importantly, they observe that the probabilities for gain, p , and loss, $1 - p$, is not enough for deciding upon whether or not to transact. The value at stake must be balanced against both p and the gain factor, i.e. the number of times the value at stake will be multiplied in case of a successful transaction. An agent is modeled by its particular risk attitude which determines for each transaction, given by the probability for gain, the gain factor and the value at stake, whether or not to accept the transaction. When the outcome of the transaction is determined by the choices of actions of another agent, the trustee, the question is to what extent this agent is trustworthy. By defining trust level as the subjective probability by which the trustee will ensure a positive outcome, a potential transaction is described by the trust level, the gain factor and the value at stake. The trustee is then trusted for the transaction in case the trust level is greater than or equal to the required probability expressed by the risk attitude.

Within the modeling by Jøsang and Presti, trust is not the determining factor in the decision making: The fundamental basis for making decisions involving risk is the individual risk attitude. For each potential transaction depending on another party, the question of trust is a question of whether this party is trusted enough.

A matter that is not accounted for in the paper by Jøsang and Presti is that trust is a subjective probability estimate. This fact has implications for the management of trust since decision making should be on a well-founded basis. Subjective estimation will potentially have a margin of error introducing difficulties that should be handled.

4. Relating Trust and Risk

We will for the purpose of this paper adapt the formulation by Jøsang et al. [14], which is based on the definition provided by Gambetta [8], and define trust as follows:

Definition (Trust): Trust (or, symmetrically, distrust) is the subjective probability by which one party (the trustor) expects that another party (the trustee) performs a given action on which its welfare depends.

The important property of trust as expressed by this definition is the aspect of *belief*, i.e. trustor's subjective probability estimation. We define the level of trust to be the believed probability, so the trust level may vary from 0 (complete distrust) to 1 (complete trust).

Trustworthiness is a notion referring to the actual probability by which the trusted party will perform as expected:

Definition (Trustworthiness): Trustworthiness is the objective probability by which the trustee performs a given action on which the welfare of the trustor depends.

Well-founded trust is the case in which the trustor knows the trustworthiness of the trustee. It can be argued that in this case the trustor does not really trust. The situation can be compared to flipping a coin: You *know* that the probability for the coin to end up heads is 0.5, you do not *trust* it to do so. We may say that in cases of well-founded trust, trust reduces to calculative risk.

We will focus on the perceived or estimated trustworthiness of a potential cooperation partner as being the basis for the trustor's decision as to whether or not to cooperate. Like Gambetta, we assume that trust is a threshold point located on the upper part of the probabilistic distribution. This is illustrated by Fig. 2. The horizontal axis is the subjective probability, i.e. the level of trust, whereas the vertical axis is the objective probability, i.e. the level of trustworthiness. The t on the horizontal axis marks the trust threshold, so there is trust whenever the subjective probability is greater than t . The remaining elements of Fig. 2 will be explained stepwise below.

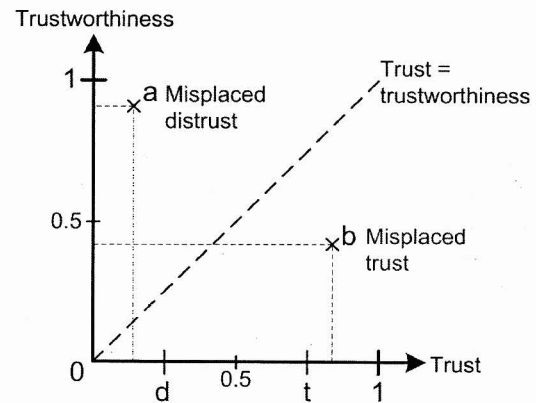


Figure 2. Trust levels

Notice, importantly, that we do not assume, as do e.g. Jøsang and Presti [15] and Marsh and Dibben [20], that the trust threshold is a cooperation threshold. It may in many situations be reasonable to avoid a cooperation with a trusted party as will be explained.

The cases in which the level of trust is less than t divides into two. In the area around 0.5 on the horizontal axis the trustee may choose any action as far as the trustor believes. She neither trusts nor distrusts him. In the lower area she believes that the trustee will act deceitfully. There is then a threshold point for *distrust*, which is marked with a d on the horizontal axis.

A crucial observation now is that a trust level $\geq t$ (resp. $\leq d$) is not enough for engaging in cooperation (resp. avoiding cooperation). As pointed out by Gambetta, the cost of deception may vary, and an extremely high cost requires a very high probability. This is where we need to understand

the exact relation between trust and risk.

Recall that a risk is defined by the probability and the consequence of an incident. The risk value is given by the function $r : P \times C \rightarrow RV$, where P is the set of probability values $[0,1]$, C is the set of consequence values and RV is the set of risk values.

In a trust relationship where p is the trust value, the subjective probability for deception is $1 - p$. Let us use the market for rubber as described by Kollock [17] as an example. At the time of sale of rubber it is impossible to tell the quality of the product, and the purchaser must determine the trustworthiness of the vendor, i.e. the probability that the vendor will provide quality goods. Assume now that $p \geq t$. Should she buy the rubber? The answer to that depends on the risk she is willing to take. Assume that the maximum risk she accepts is of value R and that she is offered to buy the rubber for the price of c . She will then buy the rubber if $R < r(1 - p, c)$. Clearly, if p is close to 1, i.e. trust is very high, the risk is low. But the risk is low also if the consequence value c closes to 0.

The trust level is hence but one of two factors that must be present in order to calculate the risk. Generally, the risk value of such a transaction is given by applying the risk function: $r(1 - \text{trust value}, \text{stake})$

What is generally missing in existing accounts on the relation between trust and risk, Jøsang and Presti [15] being an exception, is the consequence part of risk. It is not enough to consider trust only and then say that trust is risk acceptance, trust is inverse to risk, or the like. The crucial thing about trust assessment is to decide upon the upper value you are willing to invest or put at stake in a particular transaction or cooperation. More decisive than the trust threshold is then the threshold for risk acceptance. If a vendor is selling something that is suspiciously cheap, he might not be considered very trustworthy. But even if he is distrusted, he might be able to sell his products since the purchasers run low risks. The probability of the risk is extremely high, but the consequence is extremely low.

Fig. 3 illustrates the relationship between risk on the one hand and trust and stake on the other hand. This is an example in which there are three risk values, low, medium and high, each given as a function from trust and stake values. We see that in this case, the risk value is low for all trust values if the stake is close to 0. Likewise, if the stake is sufficiently high, the risk is considered high independently of the estimated trust value. The risk is generally lower when the trust value is high, but the risk value cannot be determined without considering the value at stake.

Suppose, now, that Fig. 3 models the risk attitude of a given trustor (the stakeholder). Assuming that she accepts all risks that are low, we can determine her requirements to the trustworthiness of the trustee. If, for example, the value at stake in a transaction is s_2 she must trust the trustee with

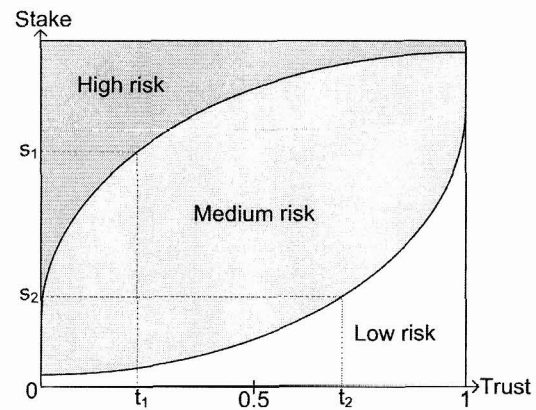


Figure 3. Risk and trust

a value $\geq t_2$. Symmetrically, if the trust value is already determined to be t_2 , she will do the transactions in which the stake is $\leq s_2$.

Assuming that the high risks are unacceptable, she will avoid all transactions in which the trust value is t_1 and the value at stake is $\geq s_1$. In the same way, if the stake is s_1 , all transactions in which the trust value is $\leq t_1$ are unacceptable.

By these observations, we see there is no need to identify trust and distrust thresholds. Not only are the notions simply irrelevant, they are also misleading if they are identified with cooperation thresholds: Someone that is extremely deceitful may then, by definition, be categorized as trusted, whereas someone that is extremely honest and truthful may, symmetrically, be categorized as distrusted.

We must not forget in all this talk about risk that there is an important positive side to trust, viz. opportunity and prospect. To place trust in someone is to run a risk, but for a purpose. To purchase rubber is to do risky business, but it also gives the opportunity to sell refined products with net profit. Recall that a prospect is defined by the probability and the (positive) consequence of an opportunity. A purchaser of rubber must define her risk acceptance level in the light of the estimated prospects involved.

Our definition of trust given above is what Jøsang et al. refer to as *reliability trust* [14]. They point out that this kind of trust is context independent in the sense that it is a measure of the perceived reliability of another party, independent of situations in which the trustor might choose to enter into a state of dependency. They observe, as we did above, that the value of what is at stake, i.e. the possible damage, must be taken into account in each specific situation. For each specific situation in a specific context there need to be present what they refer to as *decision trust*, viz. "the extent to which a given 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" [14]. Decision trust is supposed to cover elements such as utility and risk attitude. As an example they say that a person may distrust an old rope for climbing from the 3rd floor of a building during a fire exercise (reliability), but probably trust the rope in a real fire (decision).

As we see it, there is no need for the notion of decision trust. When the trustworthiness in the sense of our definition is determined, the rest is risk management. In fact, we claim that decision trust is more or less the result of risk management and that what Jøsang et al. is missing to state here is simply that trust is not always enough as a basis for the decision to cooperate. This issue is better accounted for in the paper by Jøsang and Presti [15].

5. Well-Founded Trust

Decision making is an important aspect of trust relationships, but in order to make decisions, the available information should be as accurate and complete as possible. In cases of complete information trust is well-founded, and this is where the trustor wants to be. In Fig. 2 the diagonal, dashed line marks well-foundedness, i.e. situations in which the subjective probability equals the objective probability. It is important of course that the trustor *knows* that her trust is well-founded. A trust level that is estimated on the basis of incomplete or false information may by accident equal the objective probability, but this is not well-foundedness as the aspect of ignorance is still present.

To the extent that the trustor is ignorant about the discrepancy between the believed and actual probability, there is an inconclusiveness about or a miscalculation of the involved risk. The subjective aspect of trust is thus undermining the risk estimation and management.

Fig. 2 illustrates situations in which the probability is wrongly estimated. In the area below the diagonal there is misplaced trust to various degrees: The assumed trustworthiness is higher than the actual one. The further the distance from the diagonal, the greater the misplacement. The danger of misplacing trust is not that there is a risk to it. As we have seen, risk is an intrinsic part of trust, also well-founded trust. The danger of misplacing trust is that there is more risk present than the trustor believes. Take for example the situation marked with a *b* in Fig. 2. In this case the trustor trusts the trustee since the subjective probability is greater than *t*. The trustworthiness, however, is lower than *t* and the chance of deceit is quite high.

The area above the diagonal corresponds to situations in which the assumed trustworthiness is lower than the actual one. It is perhaps less dangerous to misplace distrust than to misplace trust since it does not lead to unknown risks. Misplacing distrust should, however, be avoided since it can lead to loss of opportunities. Situation *a* in Fig. 2 is the case

in which the trustee is distrusted while actually being trustworthy with a value greater than *t*. The trustor might then unfortunately avoid cooperating with a potentially beneficial cooperation partner.

Our discussion on well-foundedness and discrepancies between subjective and objective probability relates strongly to Marsh and Dibben's discussion on the notions of trust, untrust, distrust and mistrust [20]. In our terminology their notion of trust is well-founded belief in that the trustworthiness is greater than *t*, whereas distrust is well-founded belief in that the trustworthiness is less than *d*. Untrust refers to the levels of trust below the threshold point *t*: It is not necessarily distrust, it is rather a level of trust where the trustor is more or less indifferent. We find it suitable to refer to untrust as the interval between *d* and *t* on the horizontal axis, i.e. a trust attitude that is neither very positive nor very negative. Mistrust is what we have referred to as misplaced trust. This is the case in which trust is $\geq t$ while trustworthiness is $< t$. Surprisingly Marsh and Dibben do not introduce the notion dual to mistrust, viz. misplaced distrust. It is furthermore possible to talk about misuntrust which is an estimated trust value greater than *d* and less than *t* that is either too high or too low. In cases of misuntrust the trustor will adopt an indifferent attitude towards a potential transaction partner that in reality is either deceitful or benign.

We agree with Marsh and Dibben in the importance of understanding the difference between these various notions of trust. By judging from the literature on trust management it is of particular importance to understand the implication of misplacing trust or distrust with respect to calculating risks and making decisions. Not only is the literature quite unclear about the precise relation between trust and risk, it does also not address the problem of misplacement. If trust is held to be a subjective notion, a belief, the implications of this should be dealt with.

This paper contributes both by clarifying the relation between trust and risk, emphasizing the importance of distinguishing between the subjective and objective aspects of trust relations, and by shedding new light on the notions introduced by Marsh and Dibben [20].

6. Conclusion

Risk is an intrinsic aspect of trust, but we have seen that the precise relation between the two notions has not been well understood in the literature on trust management. In this paper we have defined trust in terms of the probability that the trustee will not act deceitfully and then associated the trust value with the probability element of a risk. The decision as to whether or not to cooperate with the trustee is then not primarily determined by the estimated trustworthiness. The decisive element is the level of risk the trustor is

willing to accept as balanced against the prospects involved. The risk acceptance level is then for each potential transaction compared to both the estimated trustworthiness of the transaction partner and the value that is at stake. Transactions may hence be carried out even though the trustee is not very trustworthy.

From these observations, we see that trust is generally neither proportional nor inverse proportional to risk. Rather, as higher trustworthiness means lower probability of an incident, trust is inverse to the probability of a risk and proportional to the value the trustee is willing to stake, i.e. proportional to the consequence of a risk. The risk can hence not be determined from the trust value alone.

Risk management involves decision making, and for this to be possible the risks must be quite accurately estimated. Luhmann [11] proposes that an important aspect of trust is that it exceeds information; because of the inherent subjective aspect of trust there is generally present an inconclusiveness about the actual risk level involved within a trust relation. To avoid this inconclusiveness, trust must be well-founded. The problem, however, is not only how to *achieve* well-foundedness, it is also how to *verify* well-foundedness. A possible solution is to reduce the need for trust by replacing it with assurance, i.e. an “incentive structure that encourages benign behavior” [21].

If the ultimate goal of trust management is to enable the development of mutually beneficial cooperation [6], the focus should be on identifying and evaluating the basis for cooperation. Trust can be such a basis, however a potentially deceitful one. Ironically, then, for trust to be manageable it should be replaced with assurance.

Acknowledgment. The research on which this paper reports has partly been funded by the Research Council of Norway project ENFORCE (164382/V30). We are grateful to Cathrine Holst for valuable comments and suggestions.

References

- [1] A. Abdul-Rahman and S. Hailes. Supporting Trust in Virtual Communities. In *Proceedings of the 33rd Hawaii International Conference on System Sciences*, 2000.
- [2] AS/NZS. *Australian/New Zealand Standard, AS/NZS 4360:2004, Risk Management*, 2004.
- [3] R. Axelrod. *The Evolution of Cooperation*. Basic Books, New York, 1984.
- [4] M. Blaze, J. Feigenbaum, J. Ioannidis, and A. D. Keromytis. The Role of Trust Management in Distributed Systems Security. In *Secure Internet Programming: Security Issues for Mobile and Distributed Object*, volume 1603 of *LNCS*, pages 185–210. Springer, 1999.
- [5] M. Blaze, J. Feigenbaum, and J. Lacy. Decentralized Trust Management. In *Proceedings of the IEEE Conference on Security and Privacy*, pages 164–173, Oakland, CA, May 1996.

- [6] D. Elgesem. Normative Structures in Trust Management. In *iTrust 2006*, volume 3986 of *LNCS*, pages 48–61. Springer, 2006.
- [7] R. Falcone and C. Castelfranci. Social Trust: A Cognitive Approach. In *Trust and Deception in Virtual Societies*, pages 55–90. Kluwer Academic Publishers, 2001.
- [8] D. Gambetta. Can We Trust Trust? In *Trust: Making and Breaking Cooperative Relations*, chapter 13, pages 213–237. Department of Sociology, University of Oxford, 2000. Electronic edition.
- [9] T. Grandison and M. Sloman. A Survey of Trust in Internet Applications. *IEEE Communication Surveys*, 2000.
- [10] T. Grandison and M. Sloman. Specifying and Analysing Trust for Internet Applications. In *I3e2002*, pages 145–157, 2002.
- [11] H. Grimen. Luhmann om förtroende. Postscript in Luhmann, N.: *Förtroende. En mekanism för reduktion av social komplexitet*, 2005.
- [12] ISO/IEC. *ISO/IEC 13335, Information technology – Guidelines for management of IT security*, 1996–2000.
- [13] ISO/IEC. *ISO/IEC 17799:2000(E), Information technology – Code of practice for information security management*, 2000.
- [14] A. Jøsang, C. Keser, and T. Dimitrakos. Can We Manage Trust? In *iTrust 2005*, volume 3477 of *LNCS*, pages 93–107. Springer, 2005.
- [15] A. Jøsang and S. L. Presti. Analysing the Relationship between Risk and Trust. In *iTrust 2004*, volume 2995 of *LNCS*, pages 135–145. Springer, 2004.
- [16] A. Jøsang and N. Tran. Trust Management for E-Commerce. *Virtual Banking 2000*, 2000.
- [17] P. Kollok. The Emergence of Exchange Structures: An Experimental Study of Uncertainty, Commitment and Trust. *The American Journal of Sociology*, 100(2):313–345, 1994.
- [18] K. Konrad, G. Fuchs, and J. Barthel. Trust and Electronic Commerce – More Than a Technical Problem. In *Proceedings of the 18th Symposium in Reliable Distributed Systems*, pages 360–365. IEEE Computer Society, 1999.
- [19] N. Luhmann. Familiarity, Confidence, Trust: Problems and Alternatives. In *Trust: Making and Breaking Cooperative Relations*, chapter 6, pages 94–107. Department of Sociology, University of Oxford, 2000. Electronic edition.
- [20] S. Marsh and M. R. Dibben. Trust, Untrust, Distrust and Mistrust – An Exploration of the Dark(er) side. In *iTrust 2005*, volume 3477 of *LNCS*, pages 17–33. Springer, 2005.
- [21] L. D. Molm, N. Takahashi, and G. Peterson. Risk and Trust in Social Exchange: An Experimental Test of a Classical Proposition. *The American Journal of Sociology*, 105(5):1396–1427, 2000.
- [22] D. Povey. Developing Electronic Trust Policies Using a Risk Management Model. In *CQRE (Secure) '99*, volume 1740 of *LNCS*, pages 1–16. Springer, 1999.
- [23] J.-M. Seigneur, A. Gray, and C. D. Jensen. Trust Transfer: Encouraging Self-Recommendations without Sybil Attack. In *iTrust 2005*, volume 3477 of *LNCS*, pages 321–337. Springer, 2005.
- [24] O. E. Williamson. Calculativeness, Trust, and Economic Organization. *Journal of Law and Economics*, XXXVI:453–486, 1993.