Reasoning with Uncertain and Conflicting Opinions in Open Reputation Systems

Andreas Gutscher[1]

Institute of Communication Networks and Computer Engineering
Universität Stuttgart
Stuttgart, Germany

Abstract

Reputation systems support users to distinguish between trustworthy and malicious or unreliable services. They collect and evaluate available user opinions about services and about other users in order to determine an estimation for the trustworthiness of a specified service. The usefulness of a reputation system highly depends on its underlying trust model, i. e., the representation of trust values and the methods to calculate with these trust vales. Several proposed trust models that allow representing degrees of trust, ignorance and distrust show undesired properties when *conflicting* opinions are combined. The proposed consensus operators usually eliminate the incurred degree of conflict and perform a re-normalization. We argue that this elimination causes counterintuitive effects and should thus be avoided. Therefore, we propose a new representation of trust values that reflects also the degree of conflict, and we develop a calculus and operators to compute reputation values. Our approach requires no re-normalizations and thus avoids the thereby caused undesired effects.

Keywords: Trust model, reputation system, paraconsistent logic.

1 Introduction

The use of online services and applications (e. g., online shops, social networks and peer-to-peer applications) has become widespread in recent years. An important security problem related to online services is that users have to interact with a number of different services and other users they do not know very well and with whom they have little or no past experience. However, the users need to know whether services and other users are *trustworthy*, i. e., whether they will behave as expected (e. g., whether information sources are competent and reliable and whether services will handle disclosed personal information responsibly).

Therefore, the use of *reputation systems* has been proposed for various applications, e. g., to find reliable partners in online market places (e. g., eBay) and to detect malicious behavior in peer-to-peer and mobile ad-hoc networks. Reputation systems systematically collect available user recommendations about the trustworthiness of services and of other users, combine these statements and compute the

[1] Email: andreas.gutscher@ikr.uni-stuttgart.de

This paper is electronically published in
Electronic Notes in Theoretical Computer Science
URL: www.elsevier.nl/locate/entcs

resulting *reputation value* of all services according to a *trust model*. The trust model of a reputation system defines how to represent, reason and calculate with trust values. Designing a sound trust model is difficult because trust statements are *uncertain*. Thus, it is necessary to reason with uncertain indications and degrees of support, and it can happen that opinions *conflict* with each other.

Handling conflicting opinions has not yet been satisfactorily solved. Present trust models usually eliminate the probability mass associated with conflicting scenarios (e. g., Dempster-Shafer [11] and Jøsang [8]) or handle conflict in the same way as ignorance (Yager [12]). We argue that the degree of conflict is valuable information for the requesting application. We propose therefore a trust model that represents also the degree of conflict in the resulting reputation values. We define corresponding calculi that allow us to reason with conflicting opinions and to compute with discrete and continuous trust values.

In Sect. 2, we present related existing trust models and discuss their drawbacks. In Sect. 3, we propose our improved trust model. We demonstrate the confidence computation on an example in Sect. 4 and conclude in Sect. 5.

2 Related Work

A large number of approaches for modeling, representing, reasoning and computing with trust, reputation and uncertain information [4,11,1,8,9,10,7,12,5,6] has been proposed. In the following, we present some possibilities to represent trust relations and trust values as well as approaches to reason and compute with trust values.

2.1 Representation of Trust Relations and Trust Values

Trust is often described as the belief of a *trustor* in the competence and benevolence of a *trustee* to act honestly, reliably and dependably. Gambetta [4] describes trust as "*a particular level of the subjective probability with which an agent assesses that another agent or group of agents will perform a particular action [...] in a context in which it affects his own action*", which is more suitable in our context.

Various possibilities to represent the strength of trust relations quantitatively have been proposed. Trust values can be expressed by a fix number of discrete values (e. g., PGP/GnuPG [2]) or by continuous values, e. g., by a value $t \in [0, 1]$ (Maurer [10]). It is often beneficial to express the degree of *certainty* the trustor has in his rating in order to allow for more reliable reputation results. We therefore focus on trust models that represent degrees of trust, distrust and uncertainty. The Dempster-Shafer theory [11] is a complex mathematical theory for reasoning with uncertain evidence. In simple cases with one single proposition confidence values can be represented by a lower bound b (*belief*) and an upper bound p (*plausibility*), where $0 \leq b \leq p \leq 1$. Baldwin [1] similarly represents opinions in Fuzzy Logic with a *belief* and a *plausibility* value. In *Subjective Logic* [8] a trust value $t = (b, i, d)$ is represented by the degrees of *belief* (b), *ignorance* (i) and *disbelief* (d) (with $b, d, i \in [0..1]$, $b + d + i = 1$), which is mathematically equivalent to the previous representations via *belief* and *plausibility* ($p = b + i$, $d = 1 - p$).

[2] Pretty Good Privacy http://www.pgp.com, GNU Privacy Guard http://www.gnupg.org

2.2 Reasoning and Computing with Trust

Reputation systems collect and evaluate opinions of different entities. The collected opinions consist of trust statements with corresponding trust values and must be based on the own experience of the issuing entities (first-hand opinions). Reputation systems evaluate all available opinions from the point of view of the requesting entity according to a set of inference rules and return a computed reputation value (second-hand opinion). Inference rules define which new trust statements one can derive and how the resulting reputation values are computed from the first-hand trust values. Approaches to compute with deterministic trust values are often related to non-classical multi-valued logics (see for example [3]). We first discuss approaches with *probabilistic operators* and then approaches with *probabilistic initial views*.

2.2.1 Computation Approaches with Probabilistic Operators

Reputation systems with probabilistic operators successively merge trust statements according to the inference rules and compute the reputation value of the resulting trust statement with probabilistic operators. Unfortunately, with the proposed, non-distributive operators this approach works only in *directed series-parallel trust graphs* [9]. We nevertheless discuss proposed operators for trust values represented by *belief*, *ignorance* and *disbelief* values ($t = (b, i, d)$ with $b, i, d \in [0, 1]$ and $b+i+d = 1$). We use the symbols from Tab. 1 to present the corresponding truth tables.

Symbol	Meaning	Discrete trust value	Continuous trust value
$+$	full trust	$t' = \textit{belief}$	$t = (b, i, d) = (1, 0, 0)$
$\varnothing$	complete uncertainty	$t' = \textit{ignorance}$	$t = (b, i, d) = (0, 1, 0)$
$-$	full distrust	$t' = \textit{disbelief}$	$t = (b, i, d) = (0, 0, 1)$

Table 1
Trust values used in related work

Conjunction, Disjunction and Negation Operators

Baldwin [1] and Jøsang [8] proposed the following operators:

$$t_x \wedge t_y = \begin{pmatrix} b_x b_y \\ i_x i_y + i_x b_y + b_x i_y \\ d_x + d_y - d_x d_y \end{pmatrix}, t_x \vee t_y = \begin{pmatrix} b_x + b_y - b_x b_y \\ i_x i_y + i_x d_y + d_x i_y \\ d_x d_y \end{pmatrix}, \neg t_x = \begin{pmatrix} d_x \\ i_x \\ b_x \end{pmatrix}$$

The *belief* value b of a conjunction can be interpreted as the probability that both input values are *belief*, d as the probability that at least one input value is *disbelief*. The remaining probability mass is assigned to *ignorance*. The disjunction operator is constructed accordingly. The negation operator swaps the *belief* and *disbelief* values. From these operators we can derive the corresponding truth tables for the *discrete* trust values *belief*, *ignorance* and *disbelief* (see Fig. 1).

Recommendation Operator

A *recommendation operator* ($\otimes$) concatenates two trust relations i. e., it combines a trust relation from an entity E_A to an entity E_B with trust value t_x with a trust relation from E_B to an entity E_C with trust value t_y to one single trust relation from E_A to E_C with trust value $t_x \otimes t_y$ (see Fig. 2).

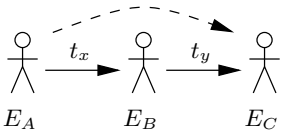
$\wedge$	$+$	$\varnothing$	$-$
$+$	$+$	$\varnothing$	$-$
$\varnothing$	$\varnothing$	$\varnothing$	$-$
$-$	$-$	$-$	$-$

$\vee$	$+$	$\varnothing$	$-$
$+$	$+$	$+$	$+$
$\varnothing$	$+$	$\varnothing$	$\varnothing$
$-$	$+$	$\varnothing$	$-$

$\neg$	
$+$	$-$
$\varnothing$	$\varnothing$
$-$	$+$

Fig. 1. Deterministic conjunction, disjunction and negation operators (Baldwin, Jøsang)

Jøsang's recommendation operator [8] follows the advice of trusted recommenders and ignores unknown and distrusted recommenders (*ignorance favoring* strategy). The operator and the corresponding truth table are shown in Fig. 2.



$$t_x \otimes t_y = \begin{pmatrix} b_x b_y \\ b_x i_y + i_x + d_x \\ b_x d_y \end{pmatrix}$$

$t'_x \otimes t'_y$		t'_x $+$	$\varnothing$	$-$
	$+$	$+$	$\varnothing$	$\varnothing$
t'_y	$\varnothing$	$\varnothing$	$\varnothing$	$\varnothing$
	$-$	$-$	$\varnothing$	$\varnothing$

Fig. 2. Recommendation operator (Jøsang)

Consensus Operator

A *consensus operator* ($\oplus$) combines the trust values of two trust relations that refer to the same proposition. Jøsang [8], Dempster-Shafer [11] and Yager [12] have proposed the consensus operators shown in Fig. 3 (undefined values are indicated by $\diamond$). Intuitively, the combination with *ignorance* does not change discrete trust

$$t_x \oplus t_y = \frac{1}{i_x + i_y - i_x i_y} \begin{pmatrix} b_x i_y + i_x b_y \\ i_x i_y \\ d_x i_y + i_x d_y \end{pmatrix}$$

$\oplus$	$+$	$\varnothing$	$-$
$+$	$\diamond$	$+$	$\diamond$
$\varnothing$	$+$	$\varnothing$	$-$
$-$	$\diamond$	$-$	$\diamond$

$$t_x \oplus t_y = \frac{1}{1 - b_x d_y - d_x b_y} \begin{pmatrix} b_x b_y + b_x i_y + i_x b_y \\ i_x i_y \\ d_x d_y + d_x i_y + i_x d_y \end{pmatrix}$$

$\oplus$	$+$	$\varnothing$	$-$
$+$	$+$	$+$	$\diamond$
$\varnothing$	$+$	$\varnothing$	$-$
$-$	$\diamond$	$-$	$-$

$$t_x \oplus t_y = \begin{pmatrix} b_x b_y + b_x i_y + i_x b_y \\ i_x i_y + b_x d_y + d_x b_y \\ d_x d_y + d_x i_y + i_x d_y \end{pmatrix}$$

$\oplus$	$+$	$\varnothing$	$-$
$+$	$+$	$+$	$\varnothing$
$\varnothing$	$+$	$\varnothing$	$-$
$-$	$\varnothing$	$-$	$-$

Fig. 3. Jøsang's (top), Dempster-Shafer's (middle) and Yager's (bottom) consensus operators

values, and in the cases of Dempster-Shafer and Yager the combination of two identical discrete trust values t' results in t'. The operators differ in their conflict handling strategy. Dempster-Shafer's operator is defined only for $1 - b_x d_y - d_x b_y > 0$, i. e., it is undefined for the combinations of *belief* with *disbelief*. The probability mass of undefined combinations is eliminated and b, i and d are re-normalized so that $b + i + d = 1$. Jøsang's operator is defined only for $i_x + i_y - i_x i_y > 0$, i. e., it is, in addition, undefined for the combinations of two belief values and of two disbelief values, which is counterintuitive as the trust values are identical. Jøsang, too, performs a re-normalization. These re-normalizations have the effect of completely

ignoring conflict and can thus lead to counterintuitive effects [13]. Yager's consensus operator [12] assigns the probability mass of conflicting combinations to *ignorance*. This avoids the counterintuitive effects of re-normalizations, but conflict is then indistinguishable from ignorance. In security-critical applications, it can be very important to distinguish these cases and treat them differently. A high degree of conflict indicates that the trustee misbehaved in the past or that some recommenders are lying, whereas ignorance indicates merely a lack of information.

2.2.2 Computation Approaches with Probabilistic Initial Views

Other approaches (e. g., Maurer [10] and Gutscher [6]) represent trust values by a value $t \in [0, 1]$. The trust values t_j of the n first-hand trust relations ($j = 1, \ldots, n$) are interpreted as probability values in the following random experiment: Each of the n trust relations is considered *valid* with the corresponding probability t_j. Then the inference rules are repeatedly applied to the *valid* trust relations and to already derived trust relations. The resulting reputation value is defined as the probability that it is possible to infer the requested reputation relation from the valid first-hand trust relations. With this approach, it is possible to evaluate trust graphs with arbitrary topology including loops and intersecting trust paths, but it is not possible to express the degree of ignorance in the trust values. We describe an adapted version of this approach in Sect. 3.4.2 in more detail.

3 New Approach for Reasoning with Conflict

We have shown that neither the handling of conflicting opinions nor the evaluation of trust graphs has been solved satisfactorily in current trust models. Therefore, we propose new representations for trust relations and trust values that reflect the degrees of belief, ignorance, disbelief and conflict (Sect. 3.1 and Sect. 3.2), and we present an approach for reasoning and computing with these trust values. This extends the trust representation in our previous approach [6], which is contained as a special case in the new approach.

We start with a deterministic calculus for discrete trust values (Sect. 3.3) and extend it for calculating with continuous values (Sect. 3.4). Our evaluation approach handles conflicting opinions reasonably, it is free from counterintuitive effects caused by re-normalizations and it can be applied to trust graphs with arbitrary topology.

3.1 Representation of Trust Relations

A trust relation is a unidirectional relation from an entity, E_A (the *trustor*) to an other entity E_B (the *trustee*). The trust relation expresses the belief of the trustor that the trustee will behave as expected with respect to some property or context r (e. g., for "taking care of my children" or "being a good dentist").

We distinguish between two *types* of trust: *functional trust* and *recommendation trust* with a certain number of *recommendation hops*. Functional trust expresses the belief that the trustee *has* the property r (e. g., "he *is* a good dentist"), whereas recommendation trust expresses the belief that the trustee can *recommend* other entities with property r over $h \geq 1$ recommendation hops (e. g., $h = 1$ expresses

the belief, that the trustee "will recommend good dentists", $h = 2$ that he "will recommend good dentist recommenders", etc.).

The belief that a certain trust relation exists is represented by a trust statement $H = \text{Trust}(E_A, E_B, r, h)$ (where $h = 0$ indicates functional trust and $h \geq 1$ recommendation trust) with an associated trust value t (see Sect. 3.2). We assume that trust relations are in general neither symmetric (i. e., $\text{Trust}(E_A, E_B, r, h)$ does not imply $\text{Trust}(E_B, E_A, r, h)$), nor transitive (i. e., $\text{Trust}(E_A, E_B, r, h)$ and $\text{Trust}(E_B, E_C, r, h)$ does not imply $\text{Trust}(E_A, E_C, r, h)$). Whether and under which conditions trust relations can be combined to trust chains should be specified explicitly with the help of inference rules (an example is given in Sect. 4).

3.2 Representing Uncertainty and Conflict in Trust Values

We reason with uncertain information, therefore, we cannot decide whether a statement H is "true" or "false". We can only collect and evaluate indications that support or refute H. Therefore, we introduce the following propositions: H^+ denotes that there are indications that support H (e. g., own experience or opinions of trustworthy entities). Similarly, H^- denotes that there are indications that refute H. However, the inability to find indications supporting H is not an indication refuting H (i. e., it does not imply H^-), and vice versa. Likewise, the existence of indications supporting H does not exclude the possibility to find indications refuting H, either (i. e., H^+ does not exclude H^-), and vice versa. First-hand opinions normally do not contain both supporting and refuting indications for H (i. e., H^+ and H^-) at the same time, but if we combine opinions of different entities it is entirely possible to find both H^+ and H^-. The latter indicates that the entities do not agree with each other and that at least some of these indications suggest the wrong conclusion, but this is not a logical contradiction.

We introduce the *discrete* trust values *belief*, *ignorance*, *disbelief* and *conflict* to represent the four possible combinations of these propositions (see Tab. 2).

Propositions	Symbol	Discrete trust value	Continuous trust value $t = (b, i, d, c)$
$\{H^+\}$	$+$	$t' = \textit{belief}$	$t = (1, 0, 0, 0)$
$\{\}$	$\varnothing$	$t' = \textit{ignorance}$	$t = (0, 1, 0, 0)$
$\{H^-\}$	$-$	$t' = \textit{disbelief}$	$t = (0, 0, 1, 0)$
$\{H^+, H^-\}$	$\pm$	$t' = \textit{conflict}$	$t = (0, 0, 0, 1)$

Table 2
Trust values in our approach

In order to describe degrees of belief, ignorance, disbelief and conflict we propose to represent *continuous* trust values by $t = (b, i, d, c)$ with $b, i, d, c \in [0, 1]$. b is the trustor's subjective estimation of the probability that there are indications supporting (but no refuting) H. Similarly, d is the subjective estimation of the probability that there are indications refuting (but no supporting) H. c is the subjective estimation of the probability that there are both supporting and refuting indications for H at the same time, and i represents the subjective estimation of the probability that there are neither supporting nor refuting indications for H. These four cases are complementary, therefore $b + i + d + c = 1$. In first-hand trust

relations c is usually 0. The correspondence between discrete and continuous trust values is shown in Tab. 2.

The trust representation in our previous approach [6] is contained as a special case in the new approach. Our previous trust value $t^* \in [0,1]$ corresponds to the degree of belief, $1-t^*$ to the degree of ignorance, hence $t = (t^*, 1-t^*, 0, 0)$.

3.3 Deterministic Operators

Inference rules define the logic of reputation systems, i. e., whether opinions can be combined and how the resulting reputation value depends on the trust values of the first-hand trust relations. In the following, we propose deterministic operators for conjunction, disjunction, negation, recommendation and consensus for the formulation of inference rules. As we favor a computation approach with probabilistic initial view it is sufficient to define these operators for discrete trust values. In Sect. 3.4, we show how these deterministic operators can be used with continuous trust values. Some examples for typical trust inference rules are shown in Sect. 4, more inference rules for reputation systems, especially also for validating the authenticity of public keys, can be found in [6].

To find the truth tables for the discrete trust values we proceed as follows: We represent the discrete trust values t'_x and t'_y of the input trust relations (H_x and H_y) as sets of propositions according to Tab. 2 (e. g., H_x^+, H_y^-). For each operator we define from which combinations of the input propositions we can infer propositions for the output reputation value (H^+, H^-). Finally, we interpret the set of output propositions as the discrete reputation value t' of the derived trust statement H.

Interestingly, this approach leads to the same truth tables for the conjunction, disjunction and negation operators as Belnap's paraconsistent four-valued logic [2] although Belnap derived these operators in a different approach from a bilattice. Belnap's logic does not provide recommendation and consensus operators though.

3.3.1 Conjunction Operator

The conjunction operator for deterministic trust values corresponds to the logical *AND*-operation and is denoted by $t' = t'_x \wedge t'_y$. The conjunction of trust statements in inference rules is denoted accordingly by $H_x \wedge H_y \Rightarrow H$. We can conclude that there are indications supporting H if we have supporting indications for both H_x and H_y. Similarly, we can conclude that there are indications refuting H if we have indications refuting H_x or H_y:

$$H_x^+, H_y^+ \Rightarrow H^+ \qquad H_x^- \Rightarrow H^- \qquad H_y^- \Rightarrow H^-$$

According to the procedure described in Sect. 3.3, we can now derive the truth table of the conjunction operator (see Fig. 4) from these two statements.

Note that the conjunction of *conflict* with *disbelief* results in *disbelief* because either H_x^- or H_y^- is sufficient to justify H^-. It is interesting that the conjunction of *conflict* with *ignorance* results in *disbelief*, too. *Conflict* for t'_x combined with *ignorance* for t'_y for example means that we can justify H_x^+ and H_x^-. H_x^- allows us to conclude H^-, but H_x^+ does not allow any conclusion without H_y^+, so that we obtain *disbelief*. The situation is different in the case of conjunction of *conflict* with *belief* which allows the justification of both H^+ and H^- and thus results in *conflict*.

$\wedge$	$+$	$\varnothing$	$-$	$\pm$
$+$	$+$	$\varnothing$	$-$	$\pm$
$\varnothing$	$\varnothing$	$\varnothing$	$-$	$-$
$-$	$-$	$-$	$-$	$-$
$\pm$	$\pm$	$-$	$-$	$\pm$

$\vee$	$+$	$\varnothing$	$-$	$\pm$
$+$	$+$	$+$	$+$	$+$
$\varnothing$	$+$	$\varnothing$	$\varnothing$	$+$
$-$	$+$	$\varnothing$	$-$	$\pm$
$\pm$	$+$	$+$	$\pm$	$\pm$

$\neg$	
$+$	$-$
$\varnothing$	$\varnothing$
$-$	$+$
$\pm$	$\pm$

Fig. 4. Our deterministic conjunction, disjunction and negation operators

3.3.2 Disjunction Operator

Similarly, the disjunction operator corresponds to the logical *OR*-operation and is denoted by $t' = t'_x \vee t'_y$. The disjunction of trust statements is denoted accordingly by $H_x \vee H_y \Rightarrow H$. We can conclude that there are indications supporting H if we have indications supporting H_x or H_y. Similarly, we can conclude that there are refuting indications for H if we have refuting indications for both H_x and H_y:

$$H_x^+ \Rightarrow H^+ \qquad H_y^+ \Rightarrow H^+ \qquad H_x^-, H_y^- \Rightarrow H^-$$

The truth table of the disjunction operator is shown in Fig. 4. The disjunction of *conflict* with *ignorance* or *belief* results in *belief* because either H_x^+ or H_y^+ is sufficient to justify H^+. It is not possible to justify H^- because this would require both H_x^- and H_y^-. The disjunction of *conflict* with *disbelief* allows the justification of both H^+ and H^- and results thus in *conflict*.

3.3.3 Negation Operator

The negation operator computes the reputation value of the opposite of a trust statement. It is denoted by $t' = \neg t'_x$ and $\neg H_x \Rightarrow H$. We can conclude that there are indications supporting H if we have indications refuting H_x, and vice versa:

$$H_x^+ \Rightarrow H^- \qquad H_x^- \Rightarrow H^+$$

The truth table of the negation operator is shown in Fig. 4.

3.3.4 Recommendation Operator

The resulting reputation values of concatenated trust relations can be calculated with the recommendation operator, which is denoted by $t' = t'_x \otimes t'_y$. The concatenation of trust statements is denoted accordingly by $H_x \otimes H_y \Rightarrow H$. We follow the *ignorance favoring* strategy of Subjective Logic [8], i. e., we ignore opinions of unknown and untrustworthy recommenders. We can thus conclude that there are indications supporting H if we have supporting indications for both H_x and H_y, and we can conclude that there are indications refuting H if we have indications supporting H_x and indications refuting H_y:

$$H_x^+, H_y^+ \Rightarrow H^+ \qquad H_x^+, H_y^- \Rightarrow H^-$$

The truth table of the recommendation operator is shown in Fig. 5. If there are no indications supporting H_x (i. e., in the cases of *ignorance* or *disbelief*) then we cannot derive any indications about H, else (i. e., in the cases of *belief* and *conflict*) we can conclude that there are indications supporting H (or refuting H) exactly if there are indications supporting H_y (or refuting H_y respectively). Therefore, the results are identical for $t'_x = \textit{belief}$ and $t'_x = \textit{conflict}$.

$t'_x \otimes t'_y$		t'_x			
		$+$	$\varnothing$	$-$	$\pm$
t'_y	$+$	$+$	$\varnothing$	$\varnothing$	$+$
	$\varnothing$	$\varnothing$	$\varnothing$	$\varnothing$	$\varnothing$
	$-$	$-$	$\varnothing$	$\varnothing$	$-$
	$\pm$	$\pm$	$\varnothing$	$\varnothing$	$\pm$

$\oplus$	$+$	$\varnothing$	$-$	$\pm$
$+$	$+$	$+$	$\pm$	$\pm$
$\varnothing$	$+$	$\varnothing$	$-$	$\pm$
$-$	$\pm$	$-$	$-$	$\pm$
$\pm$	$\pm$	$\pm$	$\pm$	$\pm$

Fig. 5. Our deterministic recommendation and consensus operators

3.3.5 Consensus Operator

The consensus operator is used to combine the trust values of two distinct opinions (t'_x and t'_y) that refer to the identical trust statement H. It calculates the cumulative reputation value, which is denoted by $t' = t'_x \oplus t'_y$. This combination of trust statements is denoted by $H_x \oplus H_y \Rightarrow H$, but this is usually not necessary because the consensus operator is applied implicitly whenever an inference rules allows deriving an already existing trust statement. We can conclude that there are indications supporting H (or refuting H) if at least one opinion has indications supporting H (or refuting H respectively), i. e., it is sufficient to unify the two sets representing the discrete trust values. The truth table of the consensus operator is shown in Fig. 5. Combining a trust value with *ignorance* or with an identical trust value does not change the trust value. Mixing *belief* and *disbelief* results in *conflict*. Conflicting trust values remain conflicting when combined with other trust values.

3.3.6 Properties of the Discrete Operators

The conjunction, disjunction and negation operators are identical to Belnap's operators [2]. Therefore the standard classical properties hold, i. e., involution ($\neg(\neg H) = H$), commutativity ($H_1 \wedge H_2 = H_2 \wedge H_1$, $H_1 \vee H_2 = H_2 \vee H_1$), associativity ($(H_1 \wedge H_2) \wedge H_3 = H_1 \wedge (H_2 \wedge H_3)$, $(H_1 \vee H_2) \vee H_3 = H_1 \vee (H_2 \vee H_3)$), distributivity ($H_1 \wedge (H_2 \vee H_3) = (H_1 \wedge H_2) \vee (H_1 \wedge H_3)$, $H_1 \vee (H_2 \wedge H_3) = (H_1 \vee H_2) \wedge (H_1 \vee H_3)$) and the De Morgan laws ($\neg(H_1 \wedge H_2) = \neg H_1 \vee \neg H_2$, $\neg(H_1 \vee H_2) = \neg H_1 \wedge \neg H_2$).

Moreover we find that consensus is commutative ($H_1 \oplus H_2 = H_2 \oplus H_1$), consensus and recommendation are associative ($(H_1 \oplus H_2) \oplus H_3 = H_1 \oplus (H_2 \oplus H_3)$, $(H_1 \otimes H_2) \otimes H_3 = H_1 \otimes (H_2 \otimes H_3)$) and that all operators are distributive over consensus ($\neg(H_1 \oplus H_2) = \neg H_1 \oplus \neg H_2$, $H_1 \wedge (H_2 \oplus H_3) = (H_1 \wedge H_2) \oplus (H_1 \wedge H_3)$, $H_1 \vee (H_2 \oplus H_3) = (H_1 \vee H_2) \oplus (H_1 \vee H_3)$, $H_1 \otimes (H_2 \oplus H_3) = (H_1 \otimes H_2) \oplus (H_1 \otimes H_3)$, $(H_1 \oplus H_2) \otimes H_3 = (H_1 \otimes H_3) \oplus (H_2 \otimes H_3)$). The latter ensures that the resulting reputation value does not depend on the order in which the inference rules are applied. This is very important to ensure consistency in trust graphs with loops and intersecting trust paths.

3.4 Reputation Computation with Discrete and Continuous Trust Values

We first describe the reputation computation with *discrete* trust values, and propose then two approaches to compute with *continuous* trust values. Due to the discussed drawbacks of the computation approach with probabilistic operators, we choose an approach with a *probabilistic initial view*. We also define corresponding probabilistic operators and show that both approaches can be mixed.

3.4.1 Reputation Computation with Discrete Trust Values

To compute the resulting discrete reputation value t' for a requested trust statement H we first represent the discrete trust values of the available first-hand opinions as sets of propositions according to Tab. 2 (e. g., $\{H_x^+\}, \{H_y^+, H_y^-\}$). Next, we repeatedly apply all inference rules to all initial and already derived propositions until no more new propositions can be derived. We test whether it was possible to derive H^+ and H^- and interpret this result set as the discrete trust value t'.

3.4.2 Reputation Computation with Probabilistic Initial View

We propose the following approach to compute with *continuous* trust values. It is based on the following random experiment: For each of the n first-hand trust statement H_j ($j = 1, \ldots, n$) with trust value $t_j = (b_j, i_j, d_j, c_j)$ we choose a discrete trust value t'_j: With probability b_j we choose *belief*, with probability i_j *ignorance*, with probability d_j *disbelief* and with probability c_j *conflict*. Next, we compute the resulting discrete reputation value t' from the chosen discrete trust values with the deterministic approach described in Sect. 3.4.1. The resulting reputation value $t = (b, i, d, c)$ is defined as follows: b is the probability that t' is *belief*, i the probability that t' is *ignorance*, d that t' is *disbelief* and c that t' is *conflict*.

This reputation value computation can be implemented with different methods. An intuitive approach to compute the exact solution is to set up a table with all possible constellations ("*possible worlds*") of the discrete trust values t'_j. For each *possible world* the resulting discrete reputation value t' is computed according to Sect. 3.4.1. For each *possible world* we compute in addition the probability p_k that this world will occur. p_k is the product of the corresponding probabilities b_j, i_j, d_j or c_j of the n trust values. To obtain the resulting belief value b we add up the probabilities p_k of all worlds in which t' is *belief*, i is the sum of the probabilities p_k of all worlds in which t' is *ignorance*, etc. This approach is illustrated in Table 3 in Sect. 4. We expect that for most applications the average length of trust chains will be quite small and that only a small fraction of the available trust statements are relevant to compute the requested reputation value, so that the computational complexity is acceptable. In many cases approximate solutions are sufficient, so that heuristics and stochastic Monte Carlo simulations can be used, too.

3.4.3 Reputation Computation with Probabilistic Operators

Let H_x and H_y be two *independent* trust statements with corresponding trust values $t_x = (b_x, i_x, d_x, c_x)$ and $t_y = (b_y, i_y, d_y, c_y)$. If we assume $H_x \wedge H_y \Rightarrow H$, $H_x \vee H_y \Rightarrow H$, $\neg H_x \Rightarrow H$, $H_x \otimes H_y \Rightarrow H$ or $H_x \oplus H_y \Rightarrow H$ as the single derivation rule of the scenario and compute the resulting reputation value for H according to Sect. 3.4.2, we obtain the reputation values shown in Fig. 6. Computing with these formulas (*probabilistic operators*) is less complex than the possible worlds approach and can thus speed up the evaluation. Although only *series-parallel trust graphs* can be evaluated completely with the probabilistic operators alone, they are nevertheless useful because the operator-based and the possible worlds computation methods can be *mixed*: It is, for example, possible to merge independent trust statements with the probabilistic operators first, and then solve the remaining, more "complex" parts of the trust graph with the possible worlds approach.

$$t_x \wedge t_y = \begin{pmatrix} b_x b_y \\ i_x i_y + i_x b_y + b_x i_y \\ d_x + d_y - d_x d_y + c_x i_y + i_x c_y \\ c_x c_y + b_x c_y + c_x b_y \end{pmatrix}, \quad t_x \vee t_y = \begin{pmatrix} b_x + b_y - b_x b_y + c_x i_y + i_x c_y \\ i_x i_y + i_x d_y + d_x i_y \\ d_x d_y \\ c_x c_y + d_x c_y + c_x d_y \end{pmatrix}$$

$$\neg t_x = (d_x, i_x, b_x, c_x)$$

$$t_x \otimes t_y = \begin{pmatrix} (b_x + c_x) b_y \\ (b_x + c_x) i_y + i_x + d_x \\ (b_x + c_x) d_y \\ (b_x + c_x) c_y \end{pmatrix}, \qquad t_x \oplus t_y = \begin{pmatrix} b_x b_y + b_x i_y + i_x b_y \\ i_x i_y \\ d_x d_y + d_x i_y + i_x d_y \\ b_x d_y + d_x b_y + c_x + c_y - c_x c_y \end{pmatrix}$$

Fig. 6. Our probabilistic conjunction, disjunction, negation, recommendation and consensus operators

4 Example

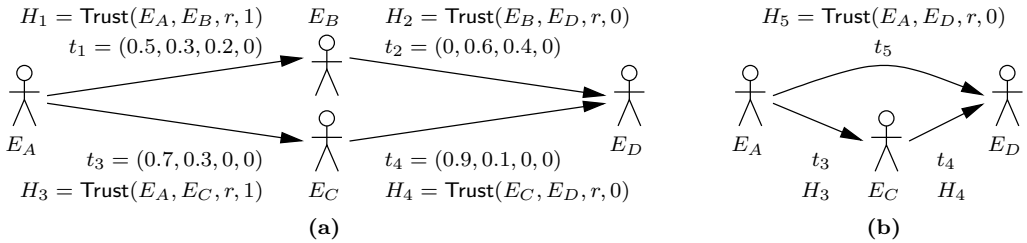


Fig. 7. Example scenario (a) and simplified example scenario (b)

In this example (Fig. 7a) we show how to compute the reputation value t for $H = \mathrm{Trust}(E_A, E_D, r, 0)$. The following inference rule[3] defines how to concatenate a recommendation trust relation ($h > 0$) with a functional trust relation:

$$\mathrm{Trust}(E_A, E_B, r, h) \otimes \mathrm{Trust}(E_B, E_C, r, 0) \Rightarrow \mathrm{Trust}(E_A, E_C, r, 0) \tag{1}$$

To demonstrate the use of the probabilistic operators (Fig. 6) we first replace the trust statements H_1 and H_2 by a new trust statement $H_5 = \mathrm{Trust}(E_A, E_D, r, 0)$ (see Fig. 7b) and compute the corresponding reputation value t_5:

$$t_5 = t_1 \otimes t_2 = (0, b_1 i_2 + i_1 + d_1, b_1 d_2, 0) = (0, 0.8, 0.2, 0)$$

Next we demonstrate the possible worlds computation approach on the resulting scenario (Fig. 7b) consisting of H_3, H_4 and H_5. In Table 3 we list all possible combinations of the discrete trust values t'_3, t'_4 and t'_5, the resulting discrete reputation value t' and the probability associated with this *world*.

t'_3	t'_4	t'_5	t'	Probability
∅	∅	−	−	$i_3 i_4 d_5 = 0.3 \cdot 0.1 \cdot 0.2 = 0.006$
∅	+	−	−	$i_3 b_4 d_5 = 0.3 \cdot 0.9 \cdot 0.2 = 0.054$
+	∅	−	−	$b_3 i_4 d_5 = 0.7 \cdot 0.1 \cdot 0.2 = 0.014$
+	+	∅	+	$b_3 b_4 i_5 = 0.7 \cdot 0.9 \cdot 0.8 = 0.504$
+	+	−	±	$b_3 b_4 d_5 = 0.7 \cdot 0.9 \cdot 0.2 = 0.126$
	else		∅	$1 - 0.006 - 0.054 - 0.014 - 0.504 - 0.126 = 0.296$

Table 3
Evaluation with possible worlds approach

[3] taken from [6], simplified and extended for reasoning with distrust

The second-to-last row for example represents the *world* with the predicates H_3^+, H_4^+ and H_5^-. H_5^- is identical to H^-. With the inference rule (1) we can conclude $H_3^+, H_4^+ \Rightarrow H^+$. The combination of the predicates H^+ and H^- results in *conflict* (implicit consensus operation). This *world* will occur with a probability of $b_3 b_4 d_5 = 0.126$. To obtain the final continuous reputation value $t = (b, i, d, c)$ we add all probabilities for each of the four discrete trust values: $b = b_3 b_4 i_5 = 0.504$, $d = i_3 i_4 d_5 + i_3 b_4 d_5 + b_3 i_4 d_5 = 0.074$, $c = b_3 b_4 d_5 = 0.126$ and $i = 1 - b - d - c = 0.296$. Thus, we obtain $t = (0.504, 0.296, 0.074, 0.126)$.

5 Summary and Conclusions

The elimination of the probability mass associated with conflict causes counterintuitive effects and should thus be avoided. The degree of conflict is important information and its interpretation can be application specific. It should therefore be returned to the requesting application. We propose new representations for discrete and continuous trust values that reflect the degree of conflict. Therefore, we do not need any re-normalization, which avoids the known counterintuitive effects. We also propose operators to formulate inference rules as well as a deterministic and a probabilistic calculus to compute with discrete and continuous trust values. We propose a computation strategy with probabilistic initial view that can be used to evaluate trust graphs with arbitrary topology. To reduce the computational complexity we propose in addition corresponding probabilistic operators and show that both approaches can be mixed in order to speed up the reputation computation.

References

[1] Baldwin, J. F., *Evidential Support Logic Programming*, Fuzzy Sets Systems **24** (1987), pp. 1–26.

[2] Belnap, N. D., *A Useful Four-valued Logic*, in: *Modern Uses of Multi-valued Logic*, 1975, pp. 8–37.

[3] Bergstra, J. A., I. Bethke and P. Rodenburg, *A Propositional Logic With 4 Values: True, False, Divergent and Meaningless* (1995), pp. 199–217.

[4] Gambetta, D., "Can We Trust Trust?" Basil Blackwell, 1988 pp. 213–237.

[5] Grandison, T. and M. Sloman, *A Survey of Trust in Internet Application*, IEEE Communications Surveys & Tutorials **3** (2000), pp. 2–16.

[6] Gutscher, A., *A Trust Model for an Open, Decentralized Reputation System*, in: *Proceedings of the Joint iTrust and PST Conferences on Privacy Trust Management and Security (IFIPTM 2007)*, 2007, pp. 285–300.

[7] Jonczy, J. and R. Haenni, *Credential Networks: a General Model for Distributed Trust and Authenticity Management*, in: *PST*, 2005, pp. 101–112.

[8] Jøsang, A., *Artificial Reasoning with Subjective Logic*, in: *Proceedings of the Second Australian Workshop on Commonsense Reasoning*, 1997.

[9] Jøsang, A., E. Gray and M. Kinateder, *Simplification and Analysis of Transitive Trust Networks* (2006), pp. 139–161.

[10] Maurer, U., *Modelling a Public-Key Infrastructure*, in: *Proc. 1996 European Symposium on Research in Computer Security (ESORICS' 96)*, Lecture Notes in Computer Science **1146** (1996), pp. 325–350.

[11] Shafer, G., "A Mathematical Theory of Evidence," Princeton Univ. Press, 1976.

[12] Yager, R. R., *On the Dempster-Shafer Framework and New Combination Rules*, Information Sciences **41** (1987), pp. 93–137.

[13] Zadeh, L. A., *Review of Books: A Mathematical Theory of Evidence*, The AI Magazine **5** (1984), pp. 81–83.