Possibilities and Limitations of Modeling Trust and Reputation

Andreas Gutscher[1], Jessica Heesen[2], and Oliver Siemoneit[2]

[1] Institute of Communication Networks and Computer Engineering, Universität Stuttgart, Germany,
andreas.gutscher@ikr.uni-stuttgart.de
[2] Institute for Philosophy, Universität Stuttgart, Germany,
{jessica.heesen,oliver.siemoneit}@philo.uni-stuttgart.de

Abstract. We all highly depend and rely on the trustworthiness of information and services provided by various parties and institutions. Reputation systems are one possibility to support individuals in distinguishing trustworthy partners from malicious and unreliable parties. In this paper, we discuss possibilities and limitations of different types of reputation systems and their underlying trust models. We address in especially the properties of trust relations, the quantification and representation of trust values as well as reasoning and computation with trust.

1 Introduction

Modern societies are characterized by a high level of differentiation and complexity. We all need and highly depend on information, services and applications provided by various parties and institutions. Unfortunately, it is seldom possible to verify on our own whether the information received is correct, whether a service is reliable or whether applications will be useful and run stable. Instead, we often have to rely on the experiences and expertise of others.

A reputation system is an approach to systematically evaluate opinions of online community members on various issues (e. g., products, services and events) and their opinions on the trustworthiness of other community members. The use of reputation systems has been proposed for various applications, for example to validate the trustworthiness of sellers and buyers in online auctions (e. g., in eBay), to detect free-riders in peer-to-peer networks and to ensure the authenticity of signature keys in a web of trust (e. g., in Pretty Good Privacy (PGP) / GNU Privacy Guard (GnuPG)).

However, by relying on recommendations from others we take a certain risk. Although some of the recommendations might be valuable to us, others might be misleading and harmful because some recommenders might have malicious intentions or not the required competence. Thus, we have to find out and to decide carefully whom we can trust. Unfortunately, we will not always be able to validate the trustworthiness of everyone providing recommendations on our own either, and we might want to look at recommendations about the trustworthiness

of the recommenders as well, and so on. Finally, we end up with a complex graph of trust relations. In order to evaluate the trust graph we have to know which conclusions we can draw and how we can compute the resulting strength of a derived trust relation.

The aim of this paper is to systematically explore and discuss possibilities and limitations of different types of reputation systems and their trust models. In Sect. 2 we give some basic definitions and describe properties of trust relations. Different possibilities to represent trust values are discussed in Sect. 3. In Sect. 4 we propose a classification for reputation systems, then we discuss different approaches to reason and calculate with trust in Sect. 5 and 6 respectively and conclude in Sect. 7.

2 Trust, Trustworthiness and Reputation

Before discussing if and how trust can be modeled and formally represented it should be clarified what the term "trust" might mean in particular. There exists a nearly unmanageable field of definitions for the term "trust" in literature. Trust has, for example, been defined as

- "the firm belief in the competence of an entity to act dependably, securely, and reliably within a specified context." [1]
- "a simplifying strategy that enables individuals to adapt to complex social environment, and thereby benefit from increased opportunities"[2, p. 38].
- "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 action (or independently of his capacity ever to be able to monitor it) and in a context in which it affects his own action." [3]

In consideration of this pluralism, trust could be defined abstractly as a *multi-relational concept* only: A *truster* trusts a *trustee* (e. g., a person, an institution or a technical system) in a certain *context*, if the truster has confidence in the *competence* and *intention* of the trustee and therefore beliefs that the trustee acts and behaves in an expected way, which does not harm the truster. We can therefore distinguish two categories of trust:

- *Competence trust*: Trust in the *capability* of a person, in an institution or in the functionality of a machine or a system.
- *Intentional trust*: Trust in the *moral integrity* (benevolence) of a person.

The *trust relation* between truster and trustee can be characterized as follows:

- *Symmetry*: Trust relations are *not symmetric* in general, i. e., if A trusts B, this does not necessarily imply that B trusts A.
- *Transitivity*: One can find contradictory opinions about whether "A trusts B" and "B trusts C" implies "A trusts C". According to our position trust is *not transitive*, because it is very well possible that A trusts B for performing certain actions, but not for giving recommendations[3]. Therefore,

[3] it is nevertheless possible to build trust chains under certain conditions (see Sect. 5)

the assumption of transitivity (e. g., in [4, 5]) can lead to counterintuitive effects.
- *Time Variance*: Trust may *change over time*, e. g., increase after successful co-operations and decrease after periods without interactions. This aspect will not be discussed further, though.

Trust is inherently related to risk and uncertainty. If everything would be predictable or perceivable, trust would not be required. The one who has confidence in someone or something often dares a possible harm: By acting someone exposes himself to a the risk of being disappointed in his expectations. Some people claim that "real" trust starts there, where no probability estimation could be given because of the lack of historical-empirical data. For them, trust should make a risk calculation dispensable so as to reduce complexity. "Real" trust would thus become relevant where no probability estimations can be given. However, in the context of reputation systems the term "trust" is used to refer to a risk estimation which helps to decide whether or not to choose a risky action. Unfortunately, trust is often based on a limited amount of experience, incomplete knowledge and questionable assumptions. Therefore, one should be aware of the degree of uncertainty of trust values. If a truster beliefs that he has not enough knowledge about the trustee or that he is not competent to decide on the trustworthiness of the trustee we talk about *ignorance*.

Often you act and you are not aware of the fact, that by acting you do at the same time trust in something or someone. Trust is often unconscious, is thus a way to reduce *complexity* [6] since you are not forced to explicitly control a situation which would absorb mental capacities and therefore produce extra complexity. If someone is asked to think about his (possibly unconscious) trust in others, to verbalize and to explicitly express his opinion about the trustworthiness of some trustee (including a *trust value* as a quantification of the degree of trustworthiness) to others, we obtain a *trust statement*. Trust statements make it possible to exchange opinions with others. If someone is considered trustworthy for issuing truthful and valuable trust statements (recommendation), then his opinions can be used to broaden one's own view, to learn from the experience of others and to come to more reliable trustworthiness estimations. Users may also exchange opinions about the trustworthiness of users for giving recommendations. This type of trust referring to the ability for giving trustworthy recommendations will be called *recommendation trust* in the following. To clarify the distinction we will call the direct, not recommending type of trust *functional trust*. Note that *trust category* and *trust type* are orthogonal classification dimensions as shown in Fig. 1.

The idea of issuing and exchanging trust statements leads to the design of *reputation systems*: Information systems that automatically and systematically gather trust statements of different issuers, accumulate and amalgamate the different subjective opinions and trust values according to the trustworthiness of their issuers in order to compute a resulting estimation of the trustworthiness of a given trustee, which may then serve as basis for decision making (see Fig. 2). This resulting opinion contains (in contrast to the previous trust opinions) not

Category / Type	Intention, Benevolence	Competence
Functional Trust	Trustee wants to safely land the plane (is not a terrorist).	Trustee knows how to safely land a plane (is a skilled pilot).
Recommendation Trust	Trustee gives reliable recommendations about whether others want to safely land the plane (can discover terrorists).	Trustee gives reliable recommendations about whether others can safely land a plane (can recommend skilled pilots).

Fig. 1. Exemplification of trust types and categories (trust context "landing a plane")

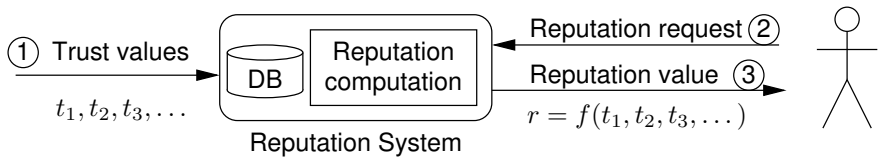


Fig. 2. Reputation System

only the opinion of one single individual, but a mixture of opinions of different individuals. To distinguish between these different types of opinions we will use the term *trust value* for the opinion of *one single entity* based only on own knowledge and experiences, whereas a *reputation value* represents a value computed from the *opinions of different entities*.

A reputation value computed by a reputation system may serve as a more reliable basis for taking decisions than the own trust value alone and can thus have an influence on the decisions taken. However, the fact that someone has a high reputation should not have a direct influence on trust values, because trust values represent the *own* opinion only without influence of the opinions of others. If a truster has only low trust in a trustee but the trustee turns out to have a high reputation, then we cannot expect (and it would not be advisable either) that the truster will somehow "increase his trust" in the trustee (how ever he would do that), but it can be advisable for the truster to engage in a risky interaction with the trustee due to the high reputation value. (If this interaction is successful, then the *own* positive experience may lead to an increase of trust, though.) Whether the truster actually does act according to the recommendation of the reputation system is not predictable, as he is not obligated to follow this recommendation. Instead, he is free to base his decision on any mixture of both, his own trust and the computed reputation value. Reputation systems therefore cannot *establish* trust between different partners of interaction, but they can convey interactions by giving the partners a broader and more reliable basis to estimate the trustworthiness of each other.

3 Modeling Trust

In order to get reputation systems work, empirical facts and circumstances need to be numerically (or symbolically) represented, i. e., the strength of trust relations has to be quantified and measured by an associated trust value. There exists a large number of proposed trust models with different approaches to represent trust values.

Besides "trust" there exist also propositions for expressing neutral or negative opinions. Although definitions have been proposed for "distrust", "untrust", "mistrust", the "lack of trust" and "ignorance" (e. g., by Marsh [7] and Grandison [1]), there is no clear consensus. One could distinguish the following forms of negative and neutral opinions: A truster *distrusts* (or *mistrusts*) a trustee if he beliefs that the trustee will not behave as expected, either due to a lack of competence or due to a malicious intention (e. g., if he beliefs that the trustee will seek to betray and actively work against him). A truster is said to have *no trust* in a trustee if he beliefs that it is neither justifiable to consider the trustee trustworthy nor to consider him distrustworthy (also called *absence of trust*).

Simple trust models represent a trust value by a single value, either on a *discrete scale*, e. g., by a Boolean value ("trust", "no trust") or by a more fine-grained scale as in PGP/GnuPG ("untrustworthy", "marginal trust" and "full trust"), or on a *continuous scale*, e. g., as proposed by Maurer [8] (trust values in the range $[0, 1]$) or by Marsh [9] (trust values in the range $[-1, 1)$).

Not all proposed trust models cover the full range of possible trust values. Some allow to express only *positive* trust values in the range between "no trust" (represented by 0) and "full trust" (represented by 1), whereas other offer the possibility to assign also "distrust" (represented by -1). However, the semantics of the trust values is sometimes different in the proposed models. Even though reasoning with distrust requires great care (an enemy of your enemy is not necessarily your friend), negative trust values may be useful especially in applications, in which the possible harm of unsuccessful interactions is high.

It is important to allow entities to express uncertainty about their trust opinions and to record this degree of uncertainty. Without this possibility the task of gathering trust opinions could cause so-called *response errors*, i. e., people who are prompted for their opinion about the trustworthiness of a subject but who do not have a reliable opinion about the trustee in question might give rather speculative answers. The degree of trustworthiness of the opinions should be taken into account in the reputation evaluation in order to avoid that valuable reliable opinions get outvoted by unreliable speculations.

Most approaches allow therefore to express ignorance (e. g., "I *can not decide* whether I can trust him") or the degree of uncertainty (e. g., "I am *quite sure* that I can judge his trustworthiness correctly") of a trust opinion, either by a discrete value (e. g., "don't know" in PGP/GnuPG) or on a continuous, additional confidence scale. Trust values can be expressed for example by two independent continuous variables for trustworthiness and confidence, e. g., trust $t \in [-1, 1]$ and uncertainty $u \in [0, 1]$ (see Fig. 3a), or by two continuous values with dependencies, e. g., Dempster-Shafer [10] and related approaches [11]

represent trust by an upper and a lower bound ($0 \leq$ belief $\leq$ plausibility ≤ 1), which is equivalent to Jøsang's opinion triangle [12] representing trust values by a belief (b), disbelief (d) and ignorance value (i) ($b, d, i \in [0, 1]$, $b + d + i = 1$, see Fig. 3b). Furthermore, it is also possible to represent trust values as arbitrary discrete distribution functions [13] (see Fig. 3c).

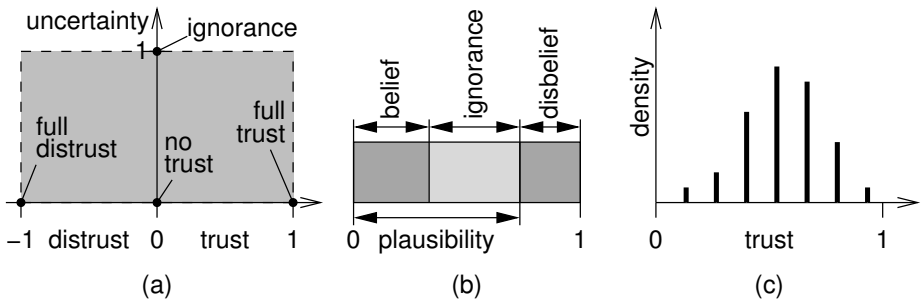


Fig. 3. Different possibilities to represent trust values

An important yet difficult task is to define the semantics of the trust values to ensure the correct interpretation of the trust statements. This may include

- defining an order relation between trust values (e. g., is "full trust" higher than "marginal trust"?),
- specifying whether differences between trust values can be meaningfully compared (e. g., is the step between "untrustworthy" and "marginal trust" comparable to the step between "marginal trust" and "full trust"?),
- specifying whether the ration between two trust values is meaningful (e. g., is "0.9" twice as trustworthy as "0.45"?), and finally
- assuring that a certain trust value (e. g., "0.45") means the same to all users (e. g., does the trust value represent a probability?).

The choice for an approach to represent reputation values may depend on the requirements and context of the application. However, approaches with the possibility to represent uncertainty make it easier to avoid counterintuitive effects. during the evaluation of trust relations.

If a truster has no information about a certain trustee, it is reasonable to assign a trust value corresponding to "ignorance" as default value. If the trust model does not allow to represent ignorance, the lowest possible trust value is a save choice to prevent malicious entities to get rid of bad reputation by changing their identity ("whitewashing").

4 Classification of Reputation Systems

We can distinguish 3 basic types of reputation systems (see Fig. 4) with different approaches to calculate reputation values:

- Type A: Flat reputation systems
- Type B: Recursively weighting reputation system
- Type C: Personalized reputation system with *trust anchor*

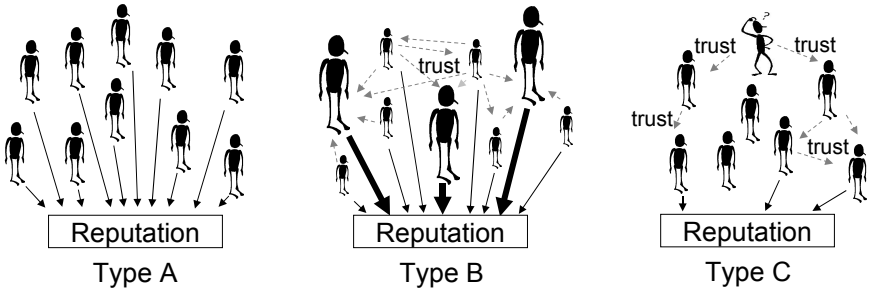


Fig. 4. Classification of reputation systems

Type A reputation systems (e. g., in eBay) are very simple. The reputation values are computed from all available trust opinions of all entities. The opinion of each entity has the same weight, i. e., liars have the same influence on the resulting trust value as honest entities. Note that (without additional measures) the collected opinions will normally not be *representative* for the group of users because the users themselves decide whether or not they want to "participate" in the "survey", i. e., to publish trust statements or not. This is especially critical if a single person can create a high number of (apparently different) entities or user accounts in a reputation system. In that case a single person can outvote all other entities by a so-called *Sybil-attack*.

Type B reputation systems (e. g., the Basic EigenTrust algorithm [14]) try to improve the quality of the computed reputation value by increasing the weight of higher ranked opinions. Reputation values of all entities are therefore computed iteratively: The new reputation values of all entities are computed from the opinions of all other entities weighted by their reputation values of the last iteration. However, the group of participating users is still not representative, and it is still possible that a large group of colluding malicious entities dominates the "public opinion" and manipulates the computed reputation values.

Type C reputation systems (e. g., as proposed by Maurer [8], Jøsang [15] and Gutscher [13]) aim to resist this kind of attacks. They always start with a "save" set of *a priori* trusted entities (the so-called *trust anchor* or *trust root*), which normally consists of the requester himself. First, only the opinions of the *a*

priori trusted entities are taken into account. Next, also the opinions of entities which have been found to be trustworthy in the previous iteration are taken into account, too. This process is repeated until the opinions of all "reachable" trustworthy entities are included in the reputation value computation. Note that opinions of untrustworthy entities are ignored as long as the opinions of the trust anchor entities are correct. In contrast to the previous reputation systems, Type C reputation systems are *personalized*, because requesters with different trust anchors will in general obtain different reputation values for the same trustee. In the following, we will focus on Type C as the most advanced type.

5 Reasoning with Trust Relations

Once the attributes, properties and the quantitative representation of trust values have been agreed upon, the process of evaluating trust relations has to be defined. For this purpose, trust models (explicitly or implicitly) define a set of inference rules, which define whether and which conclusions (new reputation relations) one can draw from a set of given trust relations. Inference rules define the made assumptions on the *transitivity property* of trust relations, but also prerequisites and restrictions depending on the type and attributes of the involved trust relations as well as on the associated trust values.

Most trust models assume that trust is *not transitive* in general, but differentiate between functional and recommendation trust and define via inference rules which trust relations can be combined to trust chains. The trust model proposed in [13] for example allows to specify for each recommendation trust relation a limit h for the allowed remaining length of trust chains (*recommendation hops*). A recommendation trust relation with $h = 1$ expresses the belief of the truster that entities recommended by the trustee are trustworthy in the sense of functional trust, whereas recommendation trust relations with $h = 2$ expresses the belief of the truster that entities recommended by other entities recommended by the trustee are trustworthy in the sense of functional trust, etc. The following trust derivation rules define how trust chains can be constructed[4]:

1. Recommendation trust from A to B with $h_1 = 1$ can be combined with functional trust from B to C to a new functional trust relation from A to C.
2. Recommendation trust from A to B with $h_1 = n + 1$ can be combined with recommendation trust from B to C with $h_2 = n$ to a new recommendation trust relation from A to C with $h = n$ (for $n \geq 1$).

These rules would allow to combine trust relations only if the number of recommendation hops matches *exactly*, which could be seen as an counterintuitive and thus inappropriate restriction. Therefore, the additional assumption was made that recommendation trust with a limit of $h = n + 1$ implies recommendation trust with a limit of $h = n$ recommendation hops (for $n \geq 1$).

[4] it is assumed that all involved trust relations refer to the same trust context

6 Computation of Reputation Values

Once new reputation relations have been derived an associated reputation value has to be computed. The computation of reputation values in Type A reputation system is very simple, e. g., the arithmetic mean all trust values is a reasonable choice. Reputation computation in Type B can be done iteratively. First, initial reputation values are computed as in Type A reputation systems. Then, new reputation values for all entities are calculated from the opinions of all entities weighted by their associated reputation values of the last iteration. This process is repeated and the reputation values will converge.

In Type C reputation systems the reputation evaluation process starts from the trust anchor specified by the requester: First, a set of all trust relations issued by *a priori* trusted entities is compiled. Then, the trust inference rules are applied to the relations in this set and all derivable trust relations are added to this set. The last step is repeated until all inferable trust relations are already contained in the set. We distinguish an *operator-based* and a *probability-theoretical* approach to compute the reputation values of the derived trust relations.

6.1 Operator-based Approach (Type C)

The trust value of the new trust relations is computed by successively combining parallel or concatenated trust relations to one single resulting trust relation (see Fig. 5). In each step, two parallel or concatenated trust relations are replaced by one resulting trust relation. The trust value of the new relation is computed from the two trust values of the replaced trust relations by a trust combination *operator*. This process is repeated until we reach a graph with one resulting trust relation from the requester to the final trustee. A simple example with operators from probability theory is shown in Fig. 5: Trust values are represented by values in the range $t \in [0, 1]$. The resulting reputation value for a concatenation of two trust relations is $t = t_1 t_2$, the resulting reputation value for parallel relations is $t = 1 - (1 - t_1)(1 - t_2) = t_1 + t_2 - t_1 t_2$. Corresponding operators handling uncertainty have been proposed for example by Dempster-Shafer [10] and related approaches [11] as well as Jøsang [12].

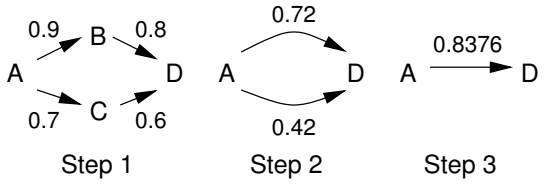


Fig. 5. Operator-based reputation computation

Reasoning with distrust requires great care to avoid possibly misleading conclusions. The following constellation is an example for a situation in which it is not obvious to decide which outcome should be considered the most "reasonable": An entity B issues a (positive or negative) trust statement about C. Entity A wants to find out whether C is trustworthy, although A distrusts B. It would be possible to *ignore* statements from untrustworthy entities or to assume the *opposite* trust value. In the first case, A might loose possibly useful information, but the strategy is a save choice. The second strategy is logically questionable (an enemy of your enemy is not necessarily your friend) and might produce misleading results, especially if B is aware of A's strategy.

The operator-based approach has a major drawback which renders it mostly unemployable: The successive combination of trust relations is only possible if either the operators are distributive (which is not true for all non-trivial operators) or if the graph of trust relations has a special structure, i. e., if it is a so-called *directed series-parallel graph* (which is unlikely to happen). A simple example of a graph which leads to this problem is shown in Fig. 6.

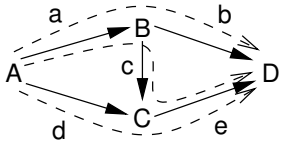
6.2 Probability Theoretical Approach (Type C)

This approach is based on the evaluation of a random experiment as proposed for example by Maurer [8] and Gutscher [13]. It is assumed that trust values are expressed by a trust value $t \in [0, 1]$, which is interpreted as the probability that the trust relation is *valid*. The resulting reputation value is the computed probability that the *requested* trust relation is *valid*, i. e., that it is possible to derive the requested trust relation from an initial *starting set*, which consists of all initially valid relations. For n initial trust relations (which can each be valid or invalid) there exist 2^n different possible starting sets. For each scenario the inference rules are applied and it is evaluated whether the desired reputation relation can be derived from the relations in the starting set. In each *successful* scenario we calculate the probability that this scenario will occur from the trust values of the initial trust relations. The resulting reputation value is the sum of the calculated probabilities of all successful scenarios.

Example: We consider the example shown in Fig. 6. The trust relations b and e represent functional trust, the trust relations c and d recommendation trust with a limit of one hop and the trust relation a recommendation trust with a limit of two hops. The corresponding trust values are $t_a, t_b, t_c, t_d, t_e \in [0, 1]$. We can find three possibilities (*trust paths*) to derive a functional reputation relation from A to D: (a, b), (d, e) or (a, c, e). The table in Fig. 6 shows for each possible starting set whether it is possible to derive a reputation relation from A to D as well as the probability of each successful scenario. The resulting reputation value r is the sum of the probabilities of all successful combinations: $r = (1 - t_a)(1 - t_b)(1 - t_c)t_d t_e + \ldots + t_a t_b t_c t_d (1 - t_e) + t_a t_b t_c t_d t_e$.

This approach can be used to evaluate arbitrary trust graphs and thus avoids the problem mentioned in Sect. 6.1, but usually has a higher computational com-

plexity[5]. Similar evaluation algorithms can be applied if the trust model supports the expression of uncertainty or if trust values are expressed by probability distributions [13].



a: Recom. trust (2 hops) t_a
b: Functional trust t_b
c: Recom. trust (1 hop) t_c
d: Recom. trust (1 hop) t_d
e: Functional trust t_e

a b c d e	A → D ?	Probability
0 0 0 0 0	no	-
0 0 0 0 1	no	-
0 0 0 1 0	no	-
0 0 0 1 1	yes	$(1-t_a)(1-t_b)(1-t_c)t_dt_e$
⋮	⋮	⋮
1 1 1 1 0	yes	$t_at_bt_ct_d(1-t_e)$
1 1 1 1 1	yes	$t_at_bt_ct_dt_e$

Fig. 6. Probability theoretical reputation computation

7 Conclusions and Outlook

We briefly presented different approaches to model, represent, reason and calculate with trust in different types of reputation systems. The main differences and limitations of reputation systems can be summarized as follows:

Type A systems are based on the undifferentiated combination of a sample of available opinions, which is *not necessarily* a representative cross-section of the population. If someone can act with a high number of (apparently different) entities, the he might outvote all other entities (Sybil-attack). Type B systems try to emphasize more *reliable* opinions by putting a higher weight on the opinions of entities which have been found "trustworthy" in the previous evaluation round. However, this does not prevent the aforementioned Sybil-attacks. Type C systems offer a promising approach to finally solve this problem by evaluating a *personal* network of trust. The evaluation process considers opinions of entities only if they are considered either trustworthy by the requester himself or if there exists a valid trust path from the requester to this entity. The proposed approaches differ in their assumptions about the transitivity property of trust and in their ways to compute the resulting reputation values. Approaches which assume trust to be *not transitive* and which distinguish functional from recommendation trust can be considered to be more precise and make it easier to avoid counterintuitive results. Operator-based reputation computation approaches usually lead to severe problems with the trust graph evaluation, which do not exist in probability theoretical approaches.

The decision for an appropriate reputation system certainly depends on the application and is often a trade-off between precision and accuracy on the one

[5] note that there exist more efficient algorithms and approximations

hand and performance and practicability on the other hand. One should be aware of the fact that the computed reputation values are always subject to uncertainty and should be used with care. Moreover there exist several inherent design problems for reputations systems which are hard to solve:

The issued trust statements usually have to be available to other users of the system. This leads to a *privacy dilemma*, because they hereby disclose very sensitive data about their personal relationships and likings. This disclosure may even lead to the effect that users give not honest but rather socially desirable ratings or act for other reasons (e. g., revenge) in a *strategic* manner. Moreover, the disclosure of trust statements could affect the interpersonal relationships of the involved users. Assigning a low trust values to someone could be interpreted as a first sign of mistrust and actually damage sensitive trust relations.

References

1. Grandison, T., Sloman, M.: A Survey of Trust in Internet Application. IEEE Communications Surveys & Tutorials **3**(4) (2000) 2–16
2. Earle, T.C., Cvetkovich, G.T.: Social Trust. Toward a Cosmopolitan Society. Praeger/Greenwood, Westport (1995)
3. Gambetta, D. In: Can We Trust Trust? Basil Blackwell (1988) 213–237 Reprinted in electronic edition from Dept. of Sociology, University of Oxford, Chapter 13.
4. Hussain, F.K., Chang, E., Dillon, T.S.: Classification of trust in Peer-to-Peer (P2P) communication. Computer Systems: Science & Engineering **19**(2) (2004)
5. Kinateder, M., Rothermel, K.: Architecture and Algorithms for a Distributed Reputation System. In: Proceedings of the First International Conference on Trust Management (iTrust 2003). Volume 2692 of LNCS. (2003) 1–16
6. Luhmann, N.: Trust: A Mechanism for the Reduction of Social Complexity. In: Trust and Power: Two Works by Niklas Luhmann. Wiley and Sons (1979)
7. Marsh, S., Dibben, M.R.: Trust, Untrust, Distrust and Mistrust – An Exploration of the Dark(er) Side. In Herrmann, P., Issarny, V., Shiu, S., eds.: Proceedings of Third iTrust International Conference (iTrust 2005), Paris, France, May 23-26, 2005. Volume 3477., Springer (May 2005) 17–33
8. Maurer, U.: Modelling a Public-Key Infrastructure. In Bertino, E., ed.: Proc. 1996 European Symposium on Research in Computer Security (ESORICS' 96). Volume 1146 of Lecture Notes in Computer Science., Springer-Verlag (1996) 325–350
9. Marsh, S.P.: Formalising Trust as a Computational Concept. PhD thesis, Department of Mathematics and Computer Science, University of Stirling (1994)
10. Shafer, G.: A Mathematical Theory of Evidence. Princeton Univ. Press (1976)
11. Sentz, K., Ferson, S.: Combination of Evidence in Dempster-Shafer Theory (2002)
12. Jøsang, A.: Artificial Reasoning with Subjective Logic. In: Proceedings of the Second Australian Workshop on Commonsense Reasoning. (1997)
13. 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) 285–300
14. Kamvar, S.D., Schlosser, M.T., Garcia-Molina, H.: The EigenTrust Algorithm for Reputation Management in P2P Networks. In: Proceedings of the 12th International Conference on World Wide Web. (2003) 640–651
15. Jøsang, A., Ismail, R.: The beta reputation system. In: Proceedings of the 15th Bled Conference on Electronic Commerce. (2002)