

# A Trust Model with Statistical Foundation

Jianqiang Shi<sup>1</sup>, Gregor v. Bochmann<sup>2</sup> and Carlisle Adams<sup>2</sup>

<sup>1</sup>*Systems Science;* <sup>2</sup>*School of Information Technology and Engineering (SITE)*

*University of Ottawa*

*Ottawa, Ontario, Canada K1N 6N5*

*{jianqshi, bochmann, cadams}@site.uottawa.ca*

**Key words:** trust, utility, decision making

**Abstract:** The widespread use of the Internet signals the need for a better understanding of trust as a basis for secure on-line interaction. In the face of increasing uncertainty and risk, users and machines must be allowed to reason effectively about the trustworthiness of other entities. In this paper, we propose a trust model that assists users and machines with decision-making in online interactions by using past behavior as a predictor of likely future behavior. We develop a general method to automatically compute trust based on self-experience and the recommendations of others. Furthermore, we apply our trust model to several utility models to increase the accuracy of decision-making in different contexts of Web Services.

## 1. INTRODUCTION

With the expansion of the Internet, users and services are often required to interact with unknown entities. This is so in application areas such as e-commerce, knowledge sharing, and even game playing. Because the entities are autonomous and potentially subject to different administrative and legal domains, it is important for each user to identify trustworthy entities or

correspondents with whom he/she should interact, and untrustworthy correspondents with whom he/she should avoid interaction [6].

Trust models have emerged as an important risk management mechanism in such online communities. The goal of a trust model is to assist users with decision-making in online interactions by using past behavior as a predictor of likely future behavior. Most electronic marketplaces on the Internet, such as eBay, Yahoo Auction, Amazon, and Epinions, support some form of trust management mechanism. eBay, for example, encourages both parties of each transaction to rate the other participant with a positive (+1), neutral (0), or negative (-1) rating. eBay makes the cumulative ratings of its members publicly known to every registered user [10]. Epinions provides a mechanism to weave “the web of trust”, a network of members whose reviews and ratings have been consistently found valuable. Each member can write a review on any topic and product. Reviews can be rated as “Not Helpful”, “Somewhat Helpful”, “Helpful”, and “Very Helpful”. The Web of Trust mimics the way people share word-of-mouth advice every day. Shareaza, a P2P file sharing system, allows members to write comments and ratings with respect to shared files. Thus, Shareaza allows members to avoid those that are fakes and download good quality, accurately represented files.

Our goal is to develop a general trust model that can be used for making rational decisions in order to make optimal choices. It should be usable in the context of Web Services and online transactions that meet real people’s needs. We have opted for a trust model that is based on stochastic models of Web Services. We will explain how trust can be built up from experimental evidence and how statistical methods can be applied, together with utility functions, to make rational choices between different service providers or different strategies for problem solving. Our trust model is scalable in the number of users and services, and is usable, both for people and artificial autonomous agents.

The rest of this paper is organized as follows. Section 2 summarizes some related work on trust models. Section 3 introduces our approach, a trust model with a statistical foundation, giving the key definitions for the state space of possible outcomes of actions, trust update, and outcome space mapping. Section 4 presents some decision models. Section 5 briefly introduces recommendations and their evaluation. Section 6 concludes the paper and discusses potential directions for future work.

## 2. EXISTING DEFINITIONS OF TRUST AND RELATED WORK

In this section we provide some existing definitions of trust and discuss some previous work related to our research.

### 2.1 Definition of trust

The trust model proposed by Marsh [1] integrates many aspects of trust taken from sociology and psychology. His model has strong sociological foundations. It is a relatively complex and abstract model and cannot be easily used in today's electronic communities. However, it is among the first works to try to give a formal treatment of trust.

Elofson [3] gives a definition suitable for our purpose. He agrees that observations are important for trust, and he defines trust as follows: "*trust is the outcome of observations leading to the belief that the actions of another may be relied upon, without explicit guarantee, to achieve a goal in a risky situation.*" Elofson notes that trust can be developed over time as the outcome of a series of confirming observations (also called the dynamics of trust).

Abdul-Rahman and Hailes [5] propose a trust model based on subjective probability, which defines four trust degrees as "Very Trustworthy", "Trustworthy", "Untrustworthy", and "Very Untrustworthy". Their aim is to provide a trust model based on the real world social properties of trust, founded on work from the social sciences. Trust is not an objective property of an agent but a subjective degree of belief about agents. A trust decision is based on the truster's relevant prior experiences and knowledge. They introduce "semantic distance" to indicate the difference between a recommendation and a truster's own perception [5].

Trust is formed and updated over time through direct interactions (self experiences) or through information (recommendations) provided by other entities of society about experiences they have had [4]. Each event that can influence the degree of trust is interpreted by the entity either as a negative or a positive experience. If the event is interpreted as a negative experience, the entity will lower his/her trust to some degree and if it is interpreted to be positive, the entity will increase trust to some degree. The change of degree of trust depends on the trust model used by the entity. The trusting entity carries out a form of continual verification and validation of the subject of trust over time [4].

## 2.2 Trust representation

One of the key issues for the design of a trust model is how trust is represented, and how the effect of experiences is specified. Representations can be qualitative, using specific qualitative labels (or term structures), or quantitative, using numbers as a representation [2].

Qualitative trust representation is sufficiently rich to specify a difference in characteristics between slow and fast dynamics, but it is not rich enough to specify more subtle differences in characteristics [2]. Such a representation causes a loss of sensitivity and accuracy. Once an entity is judged to be “Trustworthy”, then another entity that is more trustworthy, but not enough to be “Very Trustworthy”, suffers in the comparison [1].

In a quantitative representation, trust is measured by a continuous real value, which is bounded between lower and upper limits. Marsh [1] chooses to represent trust as a continuous variable over the range  $[-1, 1)$ . The value  $-1$  is the lowest possible trust value: it represents blind distrust. The value  $+1$  means blind trust. Any real number between  $-1$  and  $1$  means conditional trust. The higher the trust value, the more trustworthy the entity. However, entity-subjectivity in the use of values is a major problem. We face the problem “what does a trust of  $0.5$  mean? Is it high or low?” Such a representation of trust introduces ambiguity into the model, as the semantics of trust are usually hard to represent as a single real value [5].

Trust is a very complex and multi-faceted thing, which is hard to represent as either a single quantitative value or some qualitative rating. In either qualitative or quantitative representation, trust is a scalar, which we believe could cause potential information loss [1] or ambiguity [5]. In this paper, we propose that trust is the estimation of a probability distribution over possible outcomes of experiences. The space of possible outcomes usually depends on the context in which the trust model is used. As explained in Section 3, this approach provides a statistical foundation to the trust model and allows its application to a variety of different usage scenarios.

## 2.3 Trust recommendations

While typically, the trust that a subject has is built up by his/her direct experiences, the subject may also rely on the experiences of another subject. The latter subject is called a *recommender*, and the information provided by the recommender to the former subject is called a *trust recommendation*. The recommendation may be based on direct experiences of the recommender, and also on the trust established by the recommender, possibly through

recommendations obtained from other third parties. These issues have been discussed in several papers, such as the following.

Yu and Singh [6], focusing on acquiring information from other agents in an agent community, propose a trust model, which adapts the Dempster-Shafer Theory of Evidence to represent and propagate the ratings that agents give to their correspondents. When evaluating the trustworthiness of a correspondent, an agent combines its local evidence (based on direct prior interactions with the correspondent) with the testimonies of other agents regarding the same correspondent. Their approach includes the TrustNet representation through which the ratings can be combined. In their following work [11], Yu and Singh introduce the weighted majority technique for belief function and aggregation to detect deceptive testimonies.

Tran and Cohen [7] propose reputation-oriented reinforcement learning algorithms for buyers and sellers in open, dynamic, uncertain, and untrusted market environments. In their approach, buyers learn to avoid the risk of purchasing low quality goods and to maximize their expected value of goods by dynamically maintaining sets of reputable and disreputable sellers. Sellers learn to maximize their expected profits by adjusting product prices and optionally altering the quality of their goods. A buyer selects a seller based on his/her direct experiences without communicating with other buyers.

Chen and Singh [8] propose a hierarchical reputation structure and the confidence level of reputation. The evaluation of an individual comment is split into two parts: evaluation of its numeric rating, and of its text review. A reputation tree is built, which consists of local match, global match, and confidence. They develop a general method to compute reputations for raters based on the ratings they and others give to objects, and incorporate these reputations to generate value-added information about rated objects. The value-added information includes the reputations of raters (organized as a hierarchy of categories), the scores of objects (taking reputations of raters into account) and the confidence level of scores.

## 2.4 Different types of trust and entities involved in the trust model

Basically, we can identify three types of entities involved in a trust model: (1) the trusting entity, which we will represent by the letter  $\alpha$ , (2) the trusted entity, which we will represent by the letter  $\beta$ , and (3) possibly some recommender entities. The trusted entity is usually a kind of service provider. Typically there are several service providers and the trusting entity, a potential client, has to decide which service provider to choose for satisfying some existing needs. The purpose of the trust model is to facilitate this decision. If the trust of the entity  $\alpha$  is not well established (e.g. the entity

$\alpha$  has few direct experiences upon which to draw), the entity may rely on recommendations from recommender entities. A recommender entity could be of the same nature as the client  $\alpha$ , or it could be a recommendation server that accumulates information over time from many recommenders.

#### 2.4.1 Situational trust

A situation is a point in time relative to a specific entity. Entities find themselves in particular situations. In addition, it is the case that different situations will require different considerations with regard to trust, and will result in different values for trust, even in the same person: while I may trust Alice to drive me to school, I may not trust her to fly a plane [1]!

In this paper, we write  $T_\alpha(\beta, \delta)$  for the trust an entity  $\alpha$  has in another entity  $\beta$  with respect to a given situation  $\delta$ .

#### 2.4.2 General trust

General trust represents the trust an entity  $\alpha$  has in an entity  $\beta$ . It is not relative to any *specific* situation. It simply represents overall trust in another entity. We write  $T_\alpha(\beta)$  for the general trust of entity  $\alpha$  in entity  $\beta$ .

#### 2.4.3 Basic trust

One may ask the question: What is the initial trust an entity  $\alpha$  may have without relying on recommendations or his/her direct experiences? This initial trust reflects the trusting disposition of entity  $\alpha$ . In Marsh [1], this initial trust is called ‘basic trust’. It is derived from past experiences in all situations, indeed through the entity’s entire life time of experiences. Basic trust is not the amount of trust an entity has in any other specific entity, or situation, or environment; it is simply representative of the general trusting disposition of the entity. Considering entities as trusting entities, this allows us to simulate a basic ‘disposition’ to trust someone or something that has just been encountered. We write  $T_\alpha$  for the basic trust of entity  $\alpha$ .

### 3. A TRUST MODEL WITH A STATISTICAL FOUNDATION

In this section we propose a statistical foundation for a trust model. Such a foundation is intuitive and useful in many practical situations, as will be shown in Section 4.

### 3.1 A model of the trusted entity

Our trust model is based on a model of the trusted entity  $\beta$ . We discuss the space of possible outcomes with respect to a service performed by  $\beta$  and then propose a stochastic model for  $\beta$ .

#### 3.1.1 The space of possible outcomes

Our trust model is based on an abstract model of the trusted entity. We assume that the trust concerns the execution of a certain action by the entity. In most cases, the execution of the action corresponds to a specific service that is provided by the trusted entity. There may be different outcomes of the action. The trust is concerned with some form of prediction of what the outcome will probably be. In the case of situational trust, we are concerned with a particular action in a certain situation; in the case of general trust, the action represents any action of the trusted entity that may be of interest.

It is important to identify the space of possible outcomes. This space determines the nature of the associated trust model. We note that the granularity of this space determines the precision with which any prediction of future behavior can be made. We give in the following some typical examples.

##### a) Discrete categories

In this case, the outcomes are classified into a finite set of categories. For instance, the eBay trust model foresees the three categories: "positive", "neutral", and "negative". In the case of trust concerning the quality of the food in a restaurant, the categories may be "excellent", "good", "average", "bad", and "very bad". The case of two-valued outcomes is a special case of discrete categories; here the outcomes are classified into two categories, which may be called "good" and "bad".

While in the above examples, the different categories were ordered according to some intuitive "goodness" relationship ("good" being better than "average", for instance), there are cases in which such an ordering does not necessarily exist. We may consider the example where the outcomes are classified into the following categories: "normal: all options OK", "option A failed", and "option B failed". Here it is not clear which of the last two categories would be better.

##### b) Numeric outcomes

There are many cases in which the outcome can be characterized by a numerical value. For instance, the trust may concern the response time of a Web server, or the delivery delay of a parcel delivery service. In these cases,

we are interested in the delay for completing the action, and this delay may be measured in fractions of seconds, minutes, or hours, depending on the precision that is reasonable for the application. In these cases, the number of different outcomes is in principle infinite.

Other examples where the outcomes can be classified by a numerical value are the following: (1) What percentage of cost overruns can one expect in a construction contract? – or (2) What is the expected quality of a video obtained from a video-on-demand service?

#### c) Multidimensional outcome characterization

In many situations, the outcome of the action of interest has several parameters that are important to consider. Each of these parameters can usually be characterized either by a value from discrete value space, or a numerical (integer or real) value. In this case, we say that the space of the possible outcomes is multi-dimensional (one dimension for each parameter). Here are two examples:

1. Restaurant service with several evaluation criteria: (i) quality of food, (ii) service, and (iii) environment. For each of these three criteria, the restaurant may be classified into a certain number of discrete values, such as "excellent" down to "very bad". Therefore, the outcome of a restaurant experience may be classified as a point in this three-dimensional space, where each coordinate in this space is defined by a value between "excellent" and "very bad".
2. Multimedia presentation quality: As explained in [14] and [15], the quality of a multimedia presentation may be characterized by three values: (i) frame rate (in video frames per second), (ii) resolution (number of pixels within a frame), and (iii) color quality (number of colors distinguished per pixel). Therefore, the outcome of a video presentation obtained from a video-on-demand service may be characterized by three numerical values corresponding to these three quality of service parameters.

### 3.1.2 A stochastic model of the trusted entity

We assume that the trusted entity behaves like a stochastic process, in the sense that the outcome of an action of interest cannot be predicted exactly, that the outcome of one execution of an action of interest is statistically independent of the outcome of previous executions of that action, and that, over the long run, the probability that the outcome for the next execution of the action will be a particular point within the space of possible outcomes is described by a probability distribution, which we call the **outcome distribution** of the trusted entity, and which we represent by  $D_\beta$ . The value



of  $D_\beta$  for a particular outcome  $o \in O$  (where  $O$  is the space of possible outcomes) is written as  $D_\beta(o)$ . The outcome distribution is a distribution over the space of possible outcomes. Therefore the sum over all possible outcomes of the outcome distribution must be equal to one.

In the case of discrete outcome spaces, one usually does not make any assumptions about relationships between the outcome probabilities for different outcomes (except that they must sum to one). However, in the case of numerical outcomes, one may introduce additional assumptions. For instance, in Figure 1, a Gaussian outcome distribution is assumed, and the parameters of the Gaussian distribution are determined from a histogram of the outcomes observed during multiple experiments.



Figure 1

## 3.2 Building trust from experience

We now define trust and propose a model to build trust from prior experiences.

### 3.2.1 Definition of trust

**Definition of trust:** The trust of an entity  $\alpha$  in the outcome of an action of entity  $\beta$  is an estimation of the outcome distribution  $D_\beta$  for the execution of the action by entity  $\beta$ .

The basic mechanism for building trust is by experience, that is, by observing the execution of the action of interest by the entity  $\beta$  a certain number of times. Let us assume that the space of possible outcomes  $O$  is finite and that  $N$  observations have been made, where the outcome of the  $i$ -th observation was  $o_i$ . If we make no assumptions about relationships between outcome probabilities for different outcomes, then the best estimation of  $D_\beta$ , the trust of the observing entity  $\alpha$ , is given by the formula

$T_\alpha(\beta)(o) = (\text{number of times that the outcome } o_i \text{ was equal to } o) / N$   
(for all  $o \in O$ )

In the case that the space of possible outcomes includes a dimension with a numerical coordinate, the set of possible outcomes becomes infinite. In this case, the above simple average value calculation is not possible. Instead, the numerical coordinate is usually partitioned into a discrete number of intervals, as shown in Figure 1. Each interval is then treated like a discrete value and the above formula can be applied. If the model of the trusted entity includes an assumption about the functional form of the outcome distribution function  $D_\beta$  then the trust should be of the same form, and the parameters of this function should be adjusted to best fit the experimental data.

Instead of keeping in memory all previous experimental outcomes, one may use an incremental trust update formula. The following incremental formula is equivalent to the comprehensive formula above. For calculating the trust incrementally, we keep in memory the current trust  $T_\alpha(\beta)(o)$  for each  $o \in O$  and the number of observations to date. When a new experience yielding outcome  $o$  is observed, the values of  $T_\alpha(\beta)$  and  $N$  will be updated as follows:

$$\begin{aligned} T_\alpha(\beta)(o) &= (T_\alpha(\beta)(o) * N + 1) / (N + 1) \\ T_\alpha(\beta)(o') &= T_\alpha(\beta)(o') * N / (N + 1) \quad \text{for } o' \text{ different from } o \\ N &= N + 1 \end{aligned}$$

Note that the incremental formula and the comprehensive formula are applicable to both situational trust and general trust. In the case of the independent multidimensional outcome space,

$$T_\alpha(\beta)(o) = P(o) = \prod_{k=1}^K P(o_k) \text{ where } P(o) \text{ is the probability of outcome } o,$$

$o = (o_1, o_2, \dots, o_K)$  and  $o_k$  is the outcome in  $k$ -th dimension. Here the marginal distribution of  $o_k$  can be used instead of the joint distribution of  $o$  because the dimensions are independent.

Consider the example of restaurant service  $\beta$  whose three-dimensional outcomes are independent. Entity  $\alpha$  has situational trust in restaurant  $\beta$  based on nine experiences ( $N=9$ ) as follows

Table 1

Distribution $T(o)$	Quality of food $T1(o)$	Service $T2(o)$	Environment $T3(o)$
excellent	6/9	2/9	4/9
good	2/9	4/9	3/9
bad	1/9	2/9	1/9
very bad	0	1/9	1/9

After entity  $\alpha$  obtains one outcome such as  $o =$  (“Quality of food”= “excellent”, “service”= “good”, “environment”= “bad”), entity  $\alpha$  updates situational trust according to the incremental trust update formula and obtains the following trust ( $N=10$ )

Table 2

Distribution $T(o)$	Quality of food $T1(o)$	Service $T2(o)$	Environment $T3(o)$
excellent	7/10	2/10	4/10
good	2/10	5/10	3/10
bad	1/10	2/10	2/10
very bad	0	1/10	1/10

### 3.2.2 Estimating the error of the trust value

Generally the trusting entity estimates the true trust value with some uncertainty, both because of inherent product or service variability and because of imperfect information. Thus, it is necessary to have a method of determining the standard error of experimental outcomes. The main objective is to obtain both a desirable accuracy and a desirable confidence level with minimum cost – number of experiences.

For an outcome with a score of 0 or 1 for no or yes (Bernoulli Distribution), the standard error ( $SE$ ) of the estimated proportion  $p$ , based on random sample observations, is given by:  $SE = [p(1-p)/N]^{1/2}$  where  $p$  is the proportion obtaining a score of 1, and  $N$  is the sample size [16]. This  $SE$  is the standard deviation of the range of possible estimate values. The  $SE$  is at its maximum when  $p = 0.5$ , therefore the worst case scenario occurs when 50% are yes, and 50% are no. Under this extreme condition, the sample size,  $N$ , can then be expressed as the largest integer less than or equal to  $0.25/SE^2$ . To have some notion of the sample size, note that for  $SE$  to be 0.04 (i.e. 4%), a sample size of 156 will be needed; 5%, 100; 10%, 25.

### 3.2.3 Considering trusted entities with evolving performance

If it can be assumed that the performance of the trusted entity is not constant, but evolving over time, then the basic assumption about a given outcome distribution for the actions of the entity, valid over all times, is not

true any more. In this case, we must take into account that the outcome distribution of the trusted entity evolves over time. If the trusting entity knows the speed of this evolution, possibly defined by a given characteristic time delay, then the trusting entity may include in the trust calculation only recent experiments not older than the characteristic time delay.

It is also possible to give different weights to the different experiments, either according to their age or their order. The following incremental trust update formula based on the order of the experiments may be used:

$$T_{\alpha}(\beta)(o) = (T_{\alpha}(\beta)(o) + \gamma) / (I + \gamma)$$

$$T_{\alpha}(\beta)(o') = T_{\alpha}(\beta)(o') / (I + \gamma) \text{ for } o' \text{ different from } o$$

where the value of  $\gamma$  determines the weight of the last experience compared with the previous trust estimation.

### 3.2.4 Initial trust values

In two cases, entity  $\alpha$  needs to set his/her initial trust values in entity  $\beta$ . (i) When entities  $\alpha$  and  $\beta$  have no previous relationship (in any situation) and entity  $\alpha$  has no knowledge about entity  $\beta$ , then entity  $\alpha$  needs to initialize his/her general trust and situational trust in entity  $\beta$ . (ii) When entities  $\alpha$  and  $\beta$  have no previous relationship in a new situation but entity  $\alpha$  has general trust in entity  $\beta$ , then entity  $\alpha$  needs to initialize his/her new situational trust in entity  $\beta$ . To address these problems, a mapping between different spaces is needed. Mapping to initial trust for a particular entity or situation depends on the space of possible outcomes of that situation.

### 3.2.5 Mapping between spaces

In many cases, entity  $\alpha$  needs to map between different spaces. (i) In the case of setting initial trust values, entity  $\alpha$  needs to map his/her basic trust space to his/her general trust space, as well as his/her general trust space to his/her situational trust space. (ii) In the case of general trust update, entity  $\alpha$  needs to map his/her situational trust space to his/her general trust space. (iii) There may also be cases in which entity  $\alpha$  will update his/her basic trust as a result of a large number of general or situational trust experiences. We focus in the following two mappings:

1. Generalization mapping: from situational trust space to general trust space for the purpose of general trust update. We write  $G(o)$  for the outcome of general trust when the situational trust outcome is  $o$ . Using  $G(o)$  one can update his/her general trust  $T_{\alpha}(\beta)(G(o))$ . Note that this kind of generalization mapping causes information loss since the general trust would be more “general” (abstract) in nature and the mapping is usually a many-to-one mapping, which implies that the number of discrete

outcomes of general trust space must be no more than that of the situational trust space.

2. Specialization mapping: from basic trust space to general trust space and from general trust space to situational trust space for the purpose of setting initial trust values. We write  $S(o)$  for the outcome of situational trust when the general trust outcome is  $o$ . We also write  $S(o)$  for the outcome of general trust when the basic trust outcome is  $o$ . Using  $S(o)$  one can set initial situational trust  $T_{\alpha}(\beta, \delta)(S(o))$  and initial general trust  $T_{\alpha}(\beta)(S(o))$ . Note that the specialization mapping is the reverse process of the generalization mapping. It usually is a one-to-many mapping. An example of a mapping from the general trust to the situational trust is illustrated in the following figure.

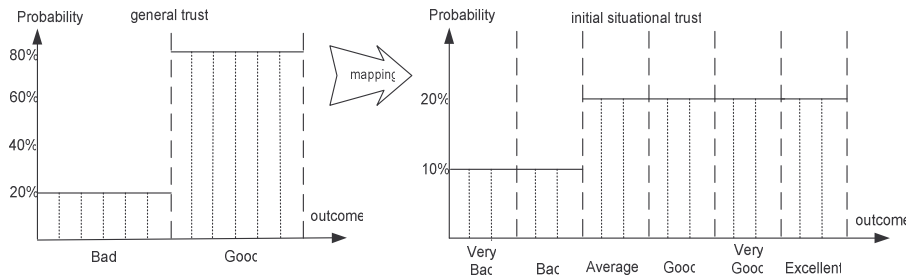


Figure 2

In this example, the entity  $\alpha$  will map his/her general trust to situational trust by defining the mapping  $S(o)$ :

- outcome “Good” in general trust maps to outcomes “average” or higher in situational trust
- outcome “Bad” in general trust maps to outcomes “Bad” or lower in situational trust

Note that the areas must be the same; that is,  $T_{\alpha}(\beta)(o) = T_{\alpha}(\beta, \delta)(S(o))$ . Thus

$$\begin{aligned}
 T_{\alpha}(\beta)(\text{“Good”}) &= T_{\alpha}(\beta, \delta)(\text{“Excellent”}) + T_{\alpha}(\beta, \delta)(\text{“Very Good”}) + \\
 &T_{\alpha}(\beta, \delta)(\text{“Good”}) + T_{\alpha}(\beta, \delta)(\text{“Average”}) = 80\%, \text{ and} \\
 T_{\alpha}(\beta)(\text{“Bad”}) &= T_{\alpha}(\beta, \delta)(\text{“Bad”}) + T_{\alpha}(\beta, \delta)(\text{“Very Bad”}) = 20\%.
 \end{aligned}$$

This histogram is then the initial set of values for situational trust outcomes (i.e.,  $N = N_{init}$ ) that will be updated over time as entity  $\alpha$  has further interactions with service  $\beta$ .

#### 4. DECISION MAKING

Decision making is often a question of selecting the optimal choice among a number of alternatives. It is therefore important to understand how different alternatives are evaluated in order to determine which is optimal. This means that for each alternative, a utility must be defined so that the alternative with the highest utility can be chosen. These kinds of approaches have been used in different areas.

In this section, we apply our trust model to several utility models to show how our trust model can be used for rational decision making. For most economic scenarios, the highest expected current utility model [13] is appropriate. For some critical scenarios, the lowest expected failure rate model [17] is appropriate. For some service scenarios, the total satisfaction model [14] is appropriate.

Expected Utility Theory (EUT) [12] states that the decision maker (DM) chooses between risky or uncertain prospects by comparing their expected utility values, i.e., the weighted sums obtained by adding the utility values of outcomes multiplied by their respective probabilities. The most popular expected utility function is the linear compensatory model in which preference for a product or service is represented by  $x_j = \sum_{k=1}^K w_k y_{jk}$  where  $x_j$

is the preference for a product or service  $j$ ,  $y_{jk}$  is the amount of attribute  $k$  in product or service  $j$ , and  $w_k$  is the importance weight assigned to attribute  $k$  [13]. In quality of service negotiation [14], a user satisfaction function plays a similar role.

Based on our trust model, we propose the following: if entity  $\alpha$  wants to use his/her trust for decision making, the entity should first establish the utility of the action of a trusted entity  $\beta$  for each possible outcome. We write  $U_\alpha(o)$  for the utility when the outcome is  $o$ . Then it is clear that the expected utility obtained from the execution of an action by entity  $\beta$  for which the trust is  $T_\alpha(\beta, \delta)$  can be calculated by the formula

$$U_\alpha(\beta) = \sum_{o \in O} T_\alpha(\beta, \delta)(o) \times U_\alpha(o)$$

In the case of multi-dimensional outcome spaces, the different dimensions may have their own utility mapping functions, and the overall utility may be the sum of the single-dimension utilities, adjusted with weight factors for the different dimensions. We then get an analogous formula to the one given in [13]. If all dimensional outcomes are independent, then the above expected utility formula can be generalized to

$$U_\alpha(\beta) = \sum_{k=1}^K U_\alpha^{(k)}(\beta) \times w_\alpha^{(k)} \text{ where } U_\alpha^{(k)}(\beta) \text{ is the expected } k\text{-th}$$

dimensional utility,  $w_{\alpha}^{(k)}$  is the subjective weight of  $k$ -th dimension (we assume that the sum of all weights is equal to 1).

We note that the latter formula corresponds to the formula for the expected utility quoted from [13] above.  $U_{\alpha}^{(k)}(\beta)$  in our formula corresponds to the value  $y_{jk}$  in the formula above.

We give three examples of making decisions and choosing the utility mapping function  $U_{\alpha}(o)$ .

1. Consider the example of restaurant service  $\beta$ . Entity  $\alpha$  assumes that all three evaluation criteria are independent. Let us assume that entity  $\alpha$  adopts the following mapping functions and dimensional weights with the following values:  $W^{(1)} = 0.6$ ;  $W^{(2)} = 0.3$ ;  $W^{(3)} = 0.1$ .

Table 3

Utility Mapping $U(o)$	Quality of food $U^{(1)}(o)$	Service $U^{(2)}(o)$	Environment $U^{(3)}(o)$
excellent	5.6	3	2
good	2.7	1	1
bad	0	-0.5	0
very bad	-4	-2	-1

The weighted “quality of food” dimension utility can be calculated using the trust values from the table in Section 3.2.1 as follows

$$U1 = \text{sum over all } o \text{ in dimension “quality of food” of } ( U^{(1)}(o) * TI(o) * W^{(1)} ) = ( 5.6 * (7/10) + 2.7 * (2/10) + 0.0 * (1/10) + (-4) * 0 ) * 0.6 = 2.676$$

Similarly, the weighted “service” dimension utility  $U2$  has the value 0.24, and the weighted “environment” dimension utility  $U3$  has the value 0.1. Therefore the utility for entity  $\alpha$  of this restaurant service  $\beta$  is  $U_{\alpha}(\beta) = U1 + U2 + U3 = 2.676 + 0.24 + 0.1 = 3.016$

Following the same process, entity  $\alpha$  can calculate the utility of other restaurant services. Entity  $\alpha$  would choose the restaurant service with the highest utility value.

2. Consider the example of multimedia presentations. Based on the multi-dimensional outcome space discussed at the end of Section 3.1.1, we could use the above formula to calculate the overall utility. However, Richards et al. [14] propose another formula. They call satisfaction  $s_k$  what we call utility  $U^{(k)}$ , and they assume that the values of satisfaction range between zero (unacceptable quality) and one (ideal quality). Instead of the weighted summation formula above, they propose to calculate the

overall satisfaction by  $S_{total} = K / \sum_{k=1}^K \frac{1}{S_k}$ . The reason for proposing this

formula is the following argument: If the satisfaction for one dimension is zero, then the total satisfaction should be zero (which is not satisfied by our formula). Both formulae satisfy the following property: If the satisfaction for all dimensions has the same value, then the overall satisfaction has that same value. Richards' formula can be extended to include weights.

3. Consider the previous example of restaurant service  $\beta$ . Entity  $\alpha$ , this time, uses a failure probability model similar to failure rate as proposed in [17] for decision making. Entity  $\alpha$  first maps the outcome space to a consideration space which consists of 2 outcomes, namely "success" and "failure"; for instance, we may assume that we have "failure" when the value of  $U_\alpha(o)$  is less than zero. The service failure probability is the proportion of outcome "failure" and can be represented by  $P_f = \sum_{U_\alpha(o) = \text{"failure"}} T_\alpha(\beta, \delta)(o)$ . The service with the lowest failure probability can be chosen. Note that one can consider this model as a special case of expected utility model in which the utility mapping has only two values, "success" and "failure".

## 5. ISSUES RELATED TO RECOMMENDATIONS

Entity  $\alpha$  can build up his/her situational and general trust from past experiences, as has been discussed in the previous sections. Due to the limitation of resources, entity  $\alpha$  may need to rely on recommendations from other entities in order to obtain trust with sufficient confidence. Entity  $\alpha$  could get many independent recommendations from different entities. Some of these recommendations will probably conflict with each other. To address the conflict, a recommendation evaluation and combination algorithm is necessary. A recommendation need not necessarily represent the real belief of the recommending entity. In fact, recommenders may lie or give out contradictory recommendations to different entities.

Following Yu and Singh [11], we define *local trust* and *global trust (reputation)*. An entity's *local trust* with respect to another entity is from his/her direct experiences. The *local trust* consists of situational trust which can be propagated to others upon request. An entity's *global trust (reputation)* with respect to another entity combines the *local trust* (if any) with recommendations received from other entities.

How to find recommenders is another issue. Yu and Singh [11] proposed an algorithm to find acyclic paths between a querying entity and recommenders. The number of possible paths is related to the connections



between entities. If the entities are densely connected, the number of paths is quite large. If the entities are sparsely connected, the number of paths could be quite small or even zero.

## 6. CONCLUSIONS AND FUTURE WORK

We have addressed the problem of building a general trust model for online entities based on their direct experiences and the recommendations of other entities. Considering trust a complex and multi-faceted thing, we use the estimated distribution in a multidimensional outcome space to represent trust. The statistical characterizations of trust (incremental trust update, estimated error, outcome space mapping) are discussed. Our trust model can be used by different decision models (utility, failure probability, satisfaction) for rational decision making in different scenarios.

For future research, we plan to investigate how the recommendations from different entities can be combined, how malicious recommendations can be detected, and how recommenders can be found. We intend to test the behavior of our trust model using simulations.

## 7. REFERENCES

- [1] Stephen Paul Marsh. *Formalising Trust as a Computational Concept*. Ph.D. Thesis, University of Stirling, April 1994
- [2] Catholijn M. Jonker, Jan Treur. *Formal Analysis of Models for the Dynamics of Trust based on Experiences*. 2<sup>nd</sup> Workshop on Deception, Fraud and Trust In Agent Societies, pp. 221 – 231, 1999
- [3] Greg Elofson. *Developing Trust with Intelligent Agents: An Exploratory Study*. In Proceedings of the 1<sup>st</sup> International Workshop on Trust, pp. 125 – 139, 1998
- [4] Miquel Montaner, Beatriz López, Josep Lluís de la Rosa. *Opinion-Based Filtering Through Trust*. In proceedings of the 6th International Workshop on Cooperative Information Agents VI, pp. 164 – 178, 2002
- [5] Alfarez Abdul-Rahman, Stephen Hailes. *Supporting Trust in Virtual Communities*. In Proceedings of the 33<sup>rd</sup> Hawaii International Conferences on System Sciences - Volume 6, 2000
- [6] Bin Yu, Munindar P. Singh. *An Evidential Model of Distributed Reputation Management*. In Proceedings of the first international joint conference on Autonomous agents and multiagent systems, pp. 294 – 301, July 2002
- [7] Thomas Tran, Robin Cohen. *Learning Algorithms for Software Agents in Uncertain and Untrusted Market Environments*. In Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence (IJCAI-03), pp. 1475 – 1476, August 2003
- [8] Mao Chen, Jaswinder Pal Singh. *Computing and Using Reputations for Internet Ratings*. EC'01, pp. 154 – 162, October 2001

- [9] Y.H. Tan and W. Thoen. *Towards a Generic Model of Trust for Electronic Commerce*. In Proceedings of the 12<sup>th</sup> International Bled Electronic Commerce Conference, Vol. 1, pp. 346 – 359, Bled Slovenia, 1999
- [10] Chrysanthos Dellarocas. *Immunizing Online Reputation Reporting Systems Against Unfair Ratings and Discriminatory Behavior*. EC'00, pp. 150 – 157, October 2000
- [11] Bin Yu, Munindar P. Singh. *Detecting Deception in Reputation management*. AAMAS'03, pp. 73 – 80, July 2003
- [12] Philippe Mongin. *Expected Utility Theory*.  
<http://expected-utility-theory.behaviouralfinance.net/Mong.pdf>
- [13] John H. Roberts, Glen L. Urban. *New Consumer Durable Brand Choice: Modeling Multiattribute Utility, Risk, and Dynamics*. Management Science Volume 34, Issue 2, 1988
- [14] Antony Richards, Glynn Rogers, Mark Antoniadis, Varuni Vitana. *Mapping User Level QoS from a Single Parameter*. In proceedings of the 2<sup>nd</sup> IFIP/IEEE International Conference on Management of Multimedia Networks and Services'98, Nov. 1998
- [15] Abdelhakim Hafid, Gregor v. Bochmann. *An Approach to Quality of Service Management in Distributed Multimedia Application: Design and an Implementation*. Multimedia Tools Appl. 9(2), 1999, pp. 167-191
- [16] Hossein Arsham. *Statistical Thinking for Managerial Decision Making*.  
<http://home.ubalt.edu/ntsbarsh/Business-stat/opre504.htm#rsss>
- [17] NIST/SEMATECH. *Engineering Statistics Handbook*  
<http://www.itl.nist.gov/div898/handbook/apr/section1/apr181.htm>