Using Evidence for Trust Computation[*]

Daniel Cvrček

Daniel.Cvrcek@cl.cam.ac.uk

Computer Laboratory
University of Cambridge
Cambridge, United Kingdom

Abstract

Trust can have its life-cycle and we can model it and utilize it for establishing secure environment for mobile environments. We assume that entities in the collaborating environment are mobile. It is not possible to perform entity enrollment. There is no globally trusted third party. Usual authentication mechanisms can not be used. We propose use of trust based on principal behaviour observations. The overall model has been devised with the SECURE project. The article makes a brief overview of the model and proposes specific approach for computation of trust values from observations. The method to be introduced is based on Dempster-Shafer theory of confirmation.

Keywords: trust, mobile, Dempster-Shafer, risk analysis.

1 Introduction

There are new challenges for information technology in the new millennium. Technologies allowing communication in global environment become actual. Mobility is another paradigm for modern communication technologies. Devices communicating through the Internet physically move on big distances and their access to the communication environment is provided by mutually independent subjects (ISPs, mobile operators, nonprofit organizations). The aim of the SECURE project [5] is to solve problems related to ubiquitous entities in mobile environment. We are forced to solve most problems in a distributed manner [2], [3], [4]. This holds also for security mechanisms like authentication and authorization.

Framework devised in the project proposes a solution based on introduction of a notion of trust. Any system offering trustworthy environment should ensure two principles.

- evaluate behaviour evaluation allows each principal to evaluate their evidence and determine trustworthiness of counterparty,
- convince other principals about my evidence and/or about trust evaluations.

There are systems that do not support mechanisms for convincing other entities about trust evaluation and it would be interesting to find limits of such systems.

We are able to describe several types of evidence for trustworthiness rating.

1. own observations
2. references
3. credentials

We have identified basic inputs and we can draw basic ideas of the model itself. The basic paradigm for the model is benefit-lost ration. We have to undergo a relatively long way to determine correct ration for particular principal and its request.

[*] The work is funded by the EU project SECURE (Secure Environment for Collaboration among Ubiquitous Roaming Entities).

Figure 1: Overall model structure

The model is based on trust so purpose of modules named *trust calculator*, *trust life-cycle manager* are obvious. We need to recognize, or more appropriately link current entity to previously obtained evidence in *entity recognition* module. Information about previous interactions with principals is in the *evidence store*, while data related to risks of particular services is in the *risk store*.

Trust data flow employs TLM module and TC (trust calculator). The resulting trust value T consists of two dimensions: trust and certainty.

T is one of the input values for real-time risk analysis. The remaining input data cover cost-PDF (probability distribution function) and parameters obtained with the service request. *Risk evaluator* combines all input data into one resulting value used for access control [1]. The following section try to solve problem of trust value T calculation.

2 Background

We are assuming only direct observations in this article and their impact on trust of an entity toward other principles. The model we are going to introduce is based on several recent works [8], [7], ...

Zhong and Bhargava try to model the time aspect in the way trust is derived from observations. They propose four basic types of modelling: equal-weight-interaction, dilute-with-time, slow-trust-quick-distrust, and slow-trust-quick-distrust-with-supervision.

Halpern and Pucella use confirmation theory and Dempster-Shafer theory of evidence to compute incremental changes of hypothesis probability. The increments are implied by functions expressing weight that a certain evidence lends to hypothesis. We can look at hypotheses as formal description of various aspects of a principal behavior.

Our target is however to determine trustworthiness of the given principal. We are going to present a method that uses several aspects of a principal behavior to determine approximate trust value for the principal.

2.1 Reasoning about Evidence

[8] presents an intuitive way of behaviour modelling. They start with a set $\mathcal{H} = \{h_1, ..., h_n\}$ of mutually exclusive and exhaustive hypotheses. It means that only one and just one hypothesis holds at a time. They also assume a finite set $\mathcal{O}$ of observations and assume that for each hypothesis h we have a probability space $(\mathcal{O}, 2^{\mathcal{O}}, \mu_h)$. It is assumed that for every observations in $\mathcal{O}$ there must be a hypothesis h such that $\mu_h(ob) > 0$. There is defined an evidence space over $\mathcal{H}$ and $\mathcal{O}$ to be a tuple $(\mathcal{H}, \mathcal{O}, \mu_{h_1}, ..., \mu_{h_n})$.

What we need is to obtain a normalized value of observation *encoding*. The authors use very simple method to compute it.

$$\omega(ob, h) = \frac{\mu_h(ob)}{\sum_{h' \in \mathcal{H}} \mu_{h'}(ob)}$$

The evidence is viewed as a *function* mapping a prior probability on the hypotheses to a posterior probability. That is,

$$\mu_{ob} = \mu_0 \oplus \omega(ob, .)$$

where the operator $\oplus$ combines two probability distributions on $\mathcal{H}$. The operator is defined by the following equation. H is a subset of hypotheses we are interested in.

$$(\mu_1 \oplus \mu_2)(H) = \frac{\sum_{h \in H} \mu_1(h)\mu_2(h)}{\sum_{h \in \mathcal{H}} \mu_1(h)\mu_2(h)} \qquad (1)$$

Finally, there is stated a theorem defining sufficient properties for weight functions ω.

Theorem: Let $\mathcal{H} = \{h_1, ..., h_m\}$ and $\mathcal{O} = \{ob_1, ..., ob_n\}$, and let f be a real-valued function with domain $\mathcal{O} \times \mathcal{H}$ such that $f(ob, h) \in [0, 1]$. Then there exists an evidence space $\mathcal{E} = (\mathcal{H}, \mathcal{O}, \mu_{h_1}, ..., \mu_{h_m})$ such that $f = \omega_{\mathcal{E}}$ if and only if f satisfies the following properties:

1. for $\forall\, ob \in \mathcal{O}$, $f(ob, .)$ is a probability measure on $\mathcal{H}$,
2. $\exists\, x_1, ..., x_n \geq 0$ such that, for all $h \in \mathcal{H}$, $\sum_{i=1}^{n} f(ob_i, f)x_i = 1$.

This theorem is important for axiomatization that is introduced in the cited article. We assume that it is not so important for our objectives but it demonstrates how an evidence space should be defined.

Example: We are going to use Black Jack as a demonstration scenario for the introduced notions and proposed method. Our reasoning about Black Jack players comprises three properties: ability to pay, willingness to pay, and play skills (let's use indexes a, w, and s for them respectively). It means that we have to define three sets of mutually exclusive and exhaustive hypotheses $\mathcal{H}_a = \{h_a^1, h_a^2, ..., h_a^{n_a}\}$, $\mathcal{H}_w = \{h_w^1, h_w^2, ..., h_w^{n_w}\}$, and $\mathcal{H}_s = \{h_s^1, h_s^2, ..., h_s^{n_s}\}$.

We need only a simple scenario, so we set $n_a = 2$ (yes/no), $n_w = 2$ (yes/no), and $n_s = 3$ (genius/scheme/looser). We can take into account e.g. following observations:

- Observations that principal has paid its debts contribute hypotheses $\omega(ob_i, h_w^{yes}$ and $\omega(ob_i, h_a^{yes})$.
- Each card played not played according to a playing scheme but winning contributes $\omega(ob_i, h_s^{genius})$.
- Each lost game sets probability of $\omega(ob_i, h_s^{genius}) = 0$.

2.2 Dynamic Trust

Other notion that we need in our constructions is dynamic trust [7]. The previous section derives new value of a hypothesis probability from previous probability and a weight of a new observations. It means that all observations have the same weight regardless on the time when they appeared. The problem was treated by Zhong and Bhargava by introducing new mapping functions for posterior probabilities. They introduced four such functions and test their behavior on four types of users (stable, repenting, cheater, and smart cheater).

Complexity of the mapping functions varies and the most complicated function is dependent on time, interaction, trustee, and it is also tunable. The proposal uses two types algorithms: trust update and trust analysis.

Trust update algorithm maintains current trust state:

$$TS_1 = f_1(Ob_1), S_i = f_i(TS_i, Ob_{i+1}),$$

Trust analysis function, however stores sequence of interactions and use it to compute new trust states. The practical ones utilize sliding window to determine which interactions to be used in computations.

$$TS_{1,i} = f_i(Ob_1, Ob_i), \;\; 1 \leq i \leq n-1$$

$$TS_{k,n} = f_n(Ob_{k-n+1}, ..., Ob_k), \;\; k \geq n$$

where $TS_{k,n}$ represents the trust state evaluated from the interaction sequence of length n starting from Ob_k.

Trust update algorithms use infinite memory model, because no evidence is discarded and its importance is just monotonically decreased. Trust analysis functions are more general. They allow us to define more complex treatment of evidence in particular history moments. This feature on the other side demands much more computation power. The authors try to find and prove existence of incremental algorithms for computation of some of their trust update algorithms.

3 Model

We assume that both approaches introduced in the previous section are very interesting but they have strong disadvantages for use in the context of SECURE project. There is no notion of time in the confirmation theory and the weight function is too simple.

The dynamic trust production does not have any mechanisms for gaining evidence or observations. It is just assumed that the values exist. The mapping function (trust update/analysis functions) try to model real situations much better regarding human understanding of trust. Their functions however do not introduce the notion of time as we use it but just as a sequence of events.

3.1 Trust and Access Rights

The SECURE framework assumes that trust value consists of two parts: information (or certainty) value and trust value. Information value expresses amount of evidence/observations that were gathered and used for the trust value computation. The trust value expresses our confidence about the principal correct behaviour.

The trust value is used to allow a particular principal access to data or functionality of our system. We can demonstrate that on categories of a user and a group commonly used in current information systems. When you operate an information system you make difference between insiders and outsiders. Insider is a user that you personally know, you know his identity. It may be your employee so there is a contract that oblige you to pay him a salary but he (principal) must abide your rules as stated in the contract. The principal is assigned a user account and a group that is associated with privileges in the information system.

The result of the SECURE framework employment should be the same. We expect the outcome to be more refined but users and roles is a good example. The framework uses two separate notions for decisions about authorization requests: risk and trust.

The risk is the notion inherent for any system. We can use a lot of methodologies to evaluate risk of the system. When we try to evaluate risks associated with users/principals in the trust based systems we come to a moment when risk and benefit of a requested authorization should be compared. This is the moment when trust comes into account. The problem of all those computations is uncertainty. The notion of uncertainty gives us yet another requirement for trust value. The value should be discrete rather than continuous because it is senseless to work with precise information in the area of trust. We have to be able to express a level of uncertainty, on the other side.

The trust as introduced in the SECURE framework is suitable for both of the requirements. The trust value should be discrete and describes e.g. a category into which a principal belongs according to evidence we have gathered. The information value can on the other hand express precisely computed fraction of satisfactory evidence we have obtained.

3.2 Basic Construction

The model we introduce assumes that we are interested in several aspects of a principal behavior. We can name financial situation, knowledge of technology, emotional characteristics to present at least a few of them. Particular evidence may contribute to just one aspect but also several different aspects of behaviour in various degree. The aspects of the behaviour are definitely not exclusive. It implies usage of several sets of hypotheses $\mathcal{H}_1, ..., \mathcal{H}_a$. Each set contains at least two hypotheses $\mathcal{H}_i = \{h_i^1, ..., h_i^{n_i}\}$.

A similar extension is necessary for observations. We can not use observations directly for evidence weight computation. We have to decide how much does an observation affect aspects of a principal behaviour. Assume that each observation has an a-tuple of aspect weights $(w_1, ..., w_a)$ with the following property:

$\forall\ w_i \in [0,1]$. The notation ob_{w_i} means value of observation ob multiplied by the aspect weight w_i. We also want to treat a real time so each observation has another attribute - time-stamp t expressed as an upper index. Now we are ready to compute set of probability values and evidence encodings.

$$\omega(ob^t_{w_i}, h^j_i) = \frac{\mu_{h^j_i}(ob^t_{w_i})}{\sum_{h'_i \in \mathcal{H}_i} \mu_{h'_i}(ob^t_{w_i})}$$

The definition of Dempster's Rule of Combination remains the same except for indexes. However, we do not use it directly for change of hypotheses probability values. We use the concept of trust update or trust analysis functions as defined in [7] instead. When defined the notion of time we can use a time function τ for adjusting weights ω.

$$\tau : O \times T \to O, \text{ or } \tau(\omega, t) \mapsto \omega$$

The parameter t (T) represents actual time (each observation has its time of occurrence as one of the attributes). The function for mapping a previous probability into a posterior one is to be defined as follows:

$$\mu^{t_{now}}_i = \mu^{t_0}_i \bigoplus_{t \in [t_0, t_{now}]} \tau(\omega(ob^t_{w_i}, h^j_i),\ t_{now})$$

That is the basic construction for obtaining probabilities for all hypotheses in any time moment. We are usually not interested in all hypotheses in real implementations. We assume that usual application will take only one hypothesis for each aspect for further computation. This approach also decreases computational requirements to determinate principal trustworthiness.

Example: We can state that only evidence not older than three months has non-zero weight. The weight of evidence back one month is 1, two months is 0.7, and three months is 0.2. Those rules define function τ.

When assuming Black Jack scenarion, we are interested only in probability of hypotheses h^{yes}_a, h^{yes}_w, h^{genius}_s, and h^{scheme}_s. That way we obtained four dimensions of *player behaviour*.

3.3 Collapse of Probabilities

The basic construction results in an a-tuple of probability values that are important for trustworthiness of a particular principal. We can understand the a-tuple as position coordinates in an a-dimensional space. This projection leads us to a simple way to collapse all the probabilities into one value - trust - as required by the SECURE framework.

The a-dimensional space, where a is determined by number of aspects important for trust computation, contains certain points representing ideal characteristics or natures of principals. When we look back to the access control we see that all needed are just definitions of groups. We assume that points in our virtual behaviour space can represent them. The necessary condition is completeness of the group set.There must be groups covering users with behaviour not allowing any access rights.

We are now prepared to compute the nearest defined ideal characteristics from the place of our principal. In fact, we have to compute distances from all ideal characteristics.

$$\delta_{aspect_i} = \sqrt{\sum_{j=1\ldots a} (\mu_j - \mu_{aspect_i})^2}$$

It shows up that the group determination is not so easy. You can easily imagine that there may occur shifts between roles that are unacceptable and we do not want them to appear. There may even be a policy that prohibits some transitions between groups.

To solve this problem we propose a solution based on a state graph: $G_t : S \to 2^S$, where $s \in S$ is an a-tuple $(\omega(h_1), \omega(h_2), ..., \omega(h_j))$ defining point in our space. The set of next states should always contain the value itself. The policy should define which roles or behaviour characteristics are incomparable and where we can define $\leq$ relation. The resulting structure can be used as a state graph defining possible transitions from one group into another one.

When we return to our computations than we have a list of distances from current user behaviour characteristics. We also know previous principal's role. When we combine all the information we get the new principal's trust value/role.

Use of such a graph brings introduces states into the model. We have to remember previous *trusts* of all principals we already know. It reminds us trust update algorithms. The new trust is the most probable value from those allowed by the state graph.

3.3.1 Certainty

The treatment of certainty part of the trustworthiness value is relatively straightforward. The policy should state a bound of the amount of evidence necessary as a total or bounds $b_{\mathcal{H}_i}$ for individual hypotheses sets. Importance $i_{\mathcal{H}_i} : \sum i_{\mathcal{H}_i} = 1$ of particular sets must be determined in the latter case. The resulting information may be got e.g from the following equation.

$$T_{certainty} = \sum_{i=1\ldots a} \frac{\iota_{\mathcal{H}_i}}{b_{\mathcal{H}_i}} i_{\mathcal{H}_i}$$

where $\iota_{\mathcal{H}_i} = \sum \tau(ob, t_{now})$ are actual amounts of evidence gathered. The equations presented are not rigorously defined but we hope they sufficiently demonstrate the idea.

4 Conclusions

We proposed a method that allows to calculate trust values from observations of principal behaviour. The method is suited for the framework designed in the SECURE project. The method calculations are relatively simple and there may be even performed certain precomputations dependent on the time function τ. We also obtain information that may be simple used for access control. We however do not know in the moment how is the method going to behave in simulations that are to take place in the following months.

There are also cons related to the model. There is a lot of configuration information that has to be set. The most difficult seems to be definition of functions for probability changes when gathering new observations. It is especially important for recommendations and credentials - information that represent aggregated evidence.

References

[1] Bacon, J., and Moody K., and Yao, W.: Access Control and Trust in the Use of Widely Distributed Services, in *Middleware 2001, Lecture Notes in Computer Science 2218*, pp. 295–310, 2001.

[2] Blaze, Matt, and Feigenbaum, Joan, and Lacy, Jack: Decentralized Trust Management,in *Proc. IEEE Conference on Security and Privacy*, AT&T, May 1996.

[3] Blaze Matt, and Feigenbaum Joan, and Keromytis, Angelos D.: The Role of Trust Management in Distributed Systems Security, in *Secure Internet Programming*, Ed. Jan Vitek and Christian Jensen, Springer-Verlag, pp. 185–210, 1999.

[4] English, Colin, and Terzis Sotirios, and Wagealla, Waleed, and Lowe, Helen, and Nixon, Paddy, and McGettrick, Andrew: Trust Dynamics in Collaborative Global Computing,in *IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE)*, June, 2003.

[5] Composite Authors: RTD Proposal - SECURE: Secure Environments for Collaboration among Ubiquitous Roaming Entities, in *project proposal*, April, 2001.

[6] Soo Hoo, Keven J.: How much is enough? A risk-management approach to computer security, *Working Paper on WWW*, Consortium for Research on Information Security and Policy, Stanford University, June 2000.

[7] Zhong Yuhui, and Bhargava, Bharat: Dynamic Trust Production Based on Interation Sequence, in *Pervasive Computing*, LNCS, Springer-Verlag, 2002.

[8] Halpern, Joseph Y. and Pucella, Riccardo: A Logic for Reasoning about Evidence, in *Journal of Artificial Intelligence Research*, pp. 57–81, No. 11,2002.