RISK BASED ACCESS CONTROL WITH UNCERTAIN AND TIME-DEPENDENT SENSITIVITY

John A. Clark[1], Juan E. Tapiador[1], John McDermid[1], Pau-Chen Cheng[2], Dakshi Agrawal[2],
Natalie Ivanic[3], Dave Slogget[4]

[1]*Deparment of Computer Science, University of York, UK. E-mail: {jac, jet, jam}@cs.york.ac.uk*
[2]*IBM Thomas J. Watson Research Center, NY, USA. E-mail: {pau, agrawal}@us.ibm.com*
[3]*US Army Research Laboratory, MD, USA. E-mail: nivanic@arl.army.mil*
[4]*LogicaCMG, UK. E-mail: dave.sloggett@logicacmg.com*

Keywords: Information Sharing; Multi-Level Security; Risk-Based Access Control.

Abstract: In traditional multi-level security (MLS) models, object labels are fixed assessments of sensitivity. In practice there will inevitably be some uncertainty about the damage that might be caused if a document falls into the wrong hands. Furthermore, unless specific management action is taken to regrade the label on an object, it does not change. This does not reflect the operational reality of many modern systems where there is clearly a temporal element to the actual sensitivity of information. Tactical information may be highly sensitive right now but comparatively irrelevant tomorrow whilst strategic secrets may need to be maintained for many years, decades, or even longer. In this paper we propose to model both security labels and clearances as probability distributions. We provide practical templates to model both uncertainty and temporally characterised dependencies, and show how these features can be naturally integrated into a recently proposed access control framework based on quantified risk.

1 Introduction

There is a recent concern about the inability of many organisations, particularly those in the national security and intelligence arena, to rapidly process, share and disseminate large quantities of sensitive information. The JASON Report (MITRE, 2004) has reinforced the view that the inflexibility of current access control models is a major inhibitor when dealing with dynamic and unpredictable environments. As an example, in the "Navy Maritime Domain Awareness Concept" paper disseminated by the US Navy in 2007 (Navy, 2007) it is recognised that non-traditional operations (e.g. when facing irregular opponents who employ asymmetric methods) generally require access to information historically unavailable to decision-makers, as well as sharing intelligence at all classification levels with other partners. Given that tasks such as sharing and disseminating information play a fundamental role in supporting informed decision making, such organisations are increasingly resorting to various ad hoc means to surpass these "cumbersome" authorisation policies (e.g., granting "temporary" authorisations for high-sensitive objects;

or, as mentioned in (MITRE, 2004), to follow the line of the old saying "it is better to ask for forgiveness rather than for permission").

An earlier paper (Chen et al., 2007a) has pointed out one major danger of such practices: they result in an unaccountable risk of information leakage. Access control is essentially about balancing risk and benefit, and a static specification of such tradeoffs is not optimal in a dynamic environment. The work in (Chen et al., 2007a) addresses this issue by making access control much more flexible. The model, known as Fuzzy MLS, is based on a *quantification* of the risk associated with every access request. Information flows are determined by particular policies, which replace the classical binary "allow/deny" decisions by a more flexible mechanism based on these risk estimators and measures of risk tolerance. Interested readers can find further details on Fuzzy MLS in (Chen et al., 2007a) and the extended version (Chen et al., 2007b).

In this paper we address two additional questions related to risk-based access control models: uncertainty and time variation of the security labels. We motivate our approach below.

1.1 Time and Sensitivity

The traditional model of multi-level security (MLS) associates security clearances with subjects, security classifications with objects, and provides a clear decision mechanism as to whether an access request should be granted or not. Thus for example, the "no read-up rule" of Bell and La Padula (BLP) model dictates that a read request should be granted only if the subject clearance dominates the object classification. The intuition behind this (and behind the corresponding "no-write down" rule) is sound. However, such rules encode for a pre-determined calculation of risks and benefits, and in many modern networking situations will preclude effective operations that can be justified on a risk basis when the specifics of the context are taken into account. Some situations demand that higher risks be taken for the sake of operational benefit. In a recent policy statement, US Director of National Intelligence Mike McConnell on 15 September 2008 said that the principal goal for risk management of any intelligence agency such as the CIA or the NSA should be to protect the agency's ability to perform its mission, *not just to protect its information assets*. One practice that certainly impedes the ability of an organisation to dispatch its responsibilities is inappropriate classification of data. The perils of underclassification are obvious; overclassification is a readily explicable outcome. But overclassification does not actually solve the problem it intends to; rather it leads to a variety of 'workarounds' and informal practices that simply take risk-based decision making outside procedural control (MITRE, 2004), effectively sweeping the issue under the carpet. Assessment of risk is an *input* into the decision making process, and it should not define the outcome under all circumstances. Closer examination of modern applications reveals further assumptions that underpin traditional MLS based access control. We shall address these in turn.

In implementations of traditional MLS models the default assumption is that the sensitivity of an object does not change over time. This principle is generally known as *tranquility* and was introduced in the BLP model to formally ensure that certain security properties hold over time[1]. For many application scenarios this clearly does not hold. In a military scenario the identified terrorist target of an air-strike is clearly

vastly more sensitive an hour before the strike than it is one hour after the strike (when the fact it has been bombed will generally be apparent to all). In contrast, the name of any pilot involved in the strike may remain sensitive for a considerable period of time. Similarly, in a commercial environment, treasury decisions on setting interest rates must be released in a controlled fashion at pre-specified times to avoid unfair market advantages. In a highly mobile tactical situation a soldier's current location may be highly sensitive, but his location yesterday will usually be less sensitive. Similar arguments hold for subject clearances. Thus, for example, a subject entering enemy territory may have his/her clearance temporarily downgraded until coming back to a safer location.

Modern collaborative operations will generate a significant amount of classified data and there would appear to be a need to prevent a general drift towards significant overclassification. More sophisticated practices will need to be adopted to ensure appropriate information usage in current times. Overclassification will make appropriate information sharing harder in almost any plausible access control scheme. Innovative risk benefit tradeoff handling approaches have been proposed to handle the inflexibility of traditional MLS, such as budget-based schemes (e.g. as suggested by (MITRE, 2004)). The price a requester pays for an access will increase with the estimate of the consequent risk, which will be inflated if the sensitivity label is too conservative. Thus, to give such innovative schemes the best chances of allowing rational risk-based decision making we must ensure that the underlying labelling accurately reflects the current sensitivity.

We clearly need also to take the time-variant nature of sensitivity into account. Traditionally this would be achieved by trusted subjects downgrading information at an appropriate time. This is a plausible approach for small numbers of documents where manual consideration can be given. However, the emergence of data-rich MANET environments forces us to reconsider this approach and ask: can we usefully model the time-varying nature of sensitivity in a principled yet practical way? In this paper we suggest some means by which this can be achieved.

1.2 Uncertain Security Levels

The traditional MLS model simply assumes that objects can be classified with an appropriate label reflecting the damage that may result from it falling into the wrong hands. There is general acceptance that such assignments are best guesses, and typically reflect the *order of magnitude* of the damage that might

---

[1]To be precise, the tranquility principle states that neither a subject clearance nor an object label must change while they are being referenced. *Strong tranquility* interprets this as that security levels does not change at all during normal operation, whilst *weak tranquility* allows changes whenever the rules of a given security policy are not violated (Bishop, 2002).

result. This is indeed a valuable construct, but in practice it will be very difficult to foresee all the implications of informational release. In particular, the value of a piece of information to an adversary may depend on what other information he has already. But in general we do not know what the enemy knows; this alone should cause us to pause and appreciate the inherent uncertainty in assigned labels. The same is applicable to the reliability of individuals (subjects). In many situations it may be impossible to assess with sufficient precision the degree of trustworthiness of a subject. Consider for example a scenario where a military operation needs the involvement of police officers and some civilians. People in these two groups ought to be provided with security clearances in order for them to have access to data. But the usual procedures employed for granting clearances in the military context (i.e., investigation of the subject's background, etc.) might simply not be affordable here.

In summary, the traditional MLS model is too strict to consider any form of uncertainty, either on the security labels or on the subjects' clearances. As pointed out in (MITRE, 2004), this limitation is particularly troublesome in multilateral and coalitional operations, where we are often required to deal with new partners in an agile, yet controlled, way. In this paper we suggest that, in principle, both security labels and users' clearances should be modelled as a probability distribution, and provide practical and plausible choices for such distributions.

1.3 Overview

In Sections 2 and 3 we provide practical templates to model time variation and uncertainty in security labels. The proposed scheme is based on the use of Beta distributions, which provides us with a suitable means to model, through parameterisation, a broad range of specifications. In Section 4 we discuss how these features can be integrated into the Fuzzy MLS access control scheme. We stress, however, that this is merely a convenient example and that the approach could be applied to other risk- or trust-based access control schemes. In Section 5 we show how the notions introduced before can be also extended to contextual information (e.g., location) considered in access-control decisions. In Section 6 we discuss how our approach relates to similar works. Finally, Section 7 concludes the paper by summarising our major contributions and pointing out some avenues for future research.

2 Modelling Uncertainty

We choose to model uncertainty in sensitivity labelling via a continuous stochastic distribution. This does not mean that sensitivities are communicated to the end users in continuous form, rather that our decision-making infrastructure uses such distributions. Sensitivity label assignment requires judgement. Some judgements will be more uncertain than others and our modelling approach must cater for such sophistications. We recall that judgements will be approximate in any case and so approximate but practical models will suffice.

Without loss of generality, we shall model sensitivity on a continuous interval $[0, S]$, $S > 0$. We make no committment to any interpretation, except that higher values correspond to higher sensitivity and vice versa. One could easily map traditional sensitivity labels onto this scale, e.g. 0 for PUBLIC, 1 for UNCLASSIFIED, 2 for RESTRICTED, 3 for CONFIDENTIAL, 4 for SECRET, 5 for TOP SECRET, etc. We would wish to allow for symmetric and skewed (both left and right) distributions, and allow different variances to be modelled. The Beta distribution provides a suitable model for our purposes. Beta distributions are defined over the interval $[0,1]$. For $\alpha, \beta > 0$, the Beta probability density function (pdf) is defined by

$$f(x; \alpha, \beta) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)} \qquad (1)$$

where $B(\alpha, \beta)$ is the beta function

$$B(x, y) = \int_0^1 t^{x-1}(1-t)^{y-1} dt \qquad (2)$$

Some useful properties of the Beta distribution along with relevant shapes are summarised in Figure 1.

We have now defined the basic Beta distributional family over the interval $[0,1]$. We extend this to the interval of interest by specifying an offset $\gamma \geq 0$ and an interval length $\lambda > 0$. Within the interval of length λ the distribution will generally be non-zero. The distribution within this interval is a stretched and normalised Beta distribution defined over $[0,1]$. Outside this interval the pdf is zero. This allows us to make statements like "the classification must be at least RESTRICTED but definitely is not TOP SECRET". This does not, of course, preclude working over the full interval range, since γ can be set to zero and λ can be equal to the full sensitivity range. The pdf can be now be defined as follows

$$g(x; \alpha, \beta, \gamma, \lambda) = \begin{cases} f(\frac{x-\gamma}{\lambda}; \alpha, \beta) & \forall x \in [\gamma, \gamma+\lambda] \\ 0 & \forall x \notin [\gamma, \gamma+\lambda] \end{cases}$$

$$(3)$$

Basic properties of the Beta distribution	
Expected value	$E(X) = \dfrac{\alpha}{\alpha + \beta}$
Variance	$Var(X) = \dfrac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$
Skewness	$\dfrac{2(\beta-\alpha)\sqrt{(\alpha+\beta+1)}}{(\alpha+\beta+2)\sqrt{(\alpha\beta)}}$
Mode	$M_0(X) = \begin{cases} \dfrac{\alpha-1}{\alpha+\beta-2} & \alpha > 1, \beta > 1 \\ 0 \text{ and } 1 & \alpha < 1, \beta < 1 \\ 0 & (\alpha < 1, \beta \geq 1) \text{ or } (\alpha = 1, \beta > 1) \\ 1 & (\alpha \geq 1, \beta < 1) \text{ or } (\alpha > 1, \beta = 1) \\ \text{Not unique} & \alpha = \beta = 1 \end{cases}$

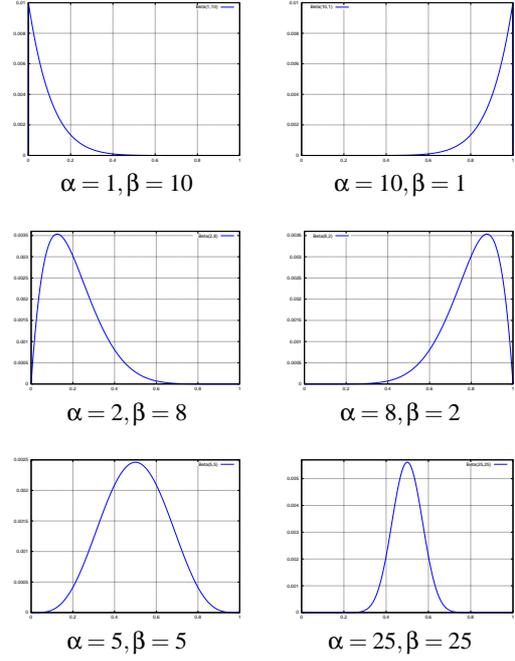| Shape |
|---|
| $\alpha = 1, \beta = 1 \Rightarrow$ Uniform [0,1] distribution |
| $\alpha < 1, \beta < 1 \Rightarrow$ U-shaped |
| $\alpha > 1, \beta > 1 \Rightarrow$ Unimodal |
| $(\alpha < 1, \beta \geq 1)$ or $(\alpha = 1, \beta > 1) \Rightarrow$ Strictly decreasing |
| $\alpha = 1, \beta \geq 2 \Rightarrow$ Convex |
| $\alpha = 1, \beta = 2 \Rightarrow$ Straight line |
| $\alpha = 1, 1 < \beta < 2 \Rightarrow$ Concave |
| $(\alpha = 1, \beta < 1)$ or $(\alpha > 1, \beta \leq 1) \Rightarrow$ Strictly increasing |
| $\alpha > 2, \beta = 1 \Rightarrow$ Convex |
| $\alpha = 2, \beta = 1 \Rightarrow$ Straight line |
| $1 < \alpha < 2, \beta = 1 \Rightarrow$ Concave |

$\alpha = 1, \beta = 10$ $\alpha = 10, \beta = 1$

$\alpha = 2, \beta = 8$ $\alpha = 8, \beta = 2$

$\alpha = 5, \beta = 5$ $\alpha = 25, \beta = 25$

Figure 1: Properties of the Beta distribution and some illustrative shapes.

Parameters α, β, γ, and λ can be chosen to provide a suitable pdf. There is some flexibility as to how such choices are made. In general γ simply shifts measures such as mean, mode, and median to the right. Parameter λ stretches and serves also to change the variance (multiplication of any random variable by a constant λ causes the variance to decrease by a factor of λ^2). The Beta distribution also allows left and right skewness to be modelled.

2.1 Mapping into Suitable Beta Distributions

The use of Beta distributions is an implementation convenience and is not intended for immediate presentation to the end users. (The average user will not take too well to being asked for parameters to a Beta distribution!) However, we can expect the user to answer plausibly question such as:

- What seems the most appropriate classification of this information (P, U, R, C, S, TS)?

- How confident are you that this classification is correct? Pretty confident or not so confident?

- Is it more likely to be classified too high than too low?

- Is there a time after which this information would cease to be classified as it is? If so, what might be the next classification?

From such questions we can provide a technical mapping to our parametric models.

An alternative to estimate the Beta parameters consists of relying on the opinions provided by a number of individuals[2]. If we assume we can collect a number of samples $x_1, \ldots, x_N$ ($N \geq 2$) regarding the sensitivity of an object, we proceed as follows. We first compute the sample mean $\bar{x}$ and variance $\bar{v}$ By using the method of moments, parameters α and β can be then estimated as

$$\alpha = \bar{x}\left(\frac{\bar{x}(1-\bar{x})}{\bar{v}} - 1\right) \qquad (4)$$

$$\beta = (1-\bar{x})\left(\frac{\bar{x}(1-\bar{x})}{\bar{v}} - 1\right) \qquad (5)$$

These estimators can be directly used to define the target distribution. If a more precise estimation is required, we can proceed iteratively as follows. Once α and β are obtained, the estimated distribution is used to generate a sufficiently large number of random samples. These are then presented to the end users, who are asked to remove values considered as definitely wrong. The resulting, "filtered" dataset is used again to produce new estimators for α and β,

---

[2]It is not the purpose of this work to provide criteria regarding how to choose such individuals. We simply assume they are personnel with appropriate qualifications to carry out such a task.

and the procedure is repeated until convergence is reached.

3 Modelling Time Variation

Above we indicated that a constant label will not reflect the true sensitivity of many aspects of data in a dynamic network environment. The true sensivity of data will exhibit some trajectory. Sensitivity may go down but, in principle, also up. Furthermore, the particular type of trajectory followed will vary with context. But if we are to handle time-variant sensitivity, we must be able to model it in some way that the user accepts as plausibly reflecting operational reality. In this section we provide a variety of simple templates for temporal dependencies of sensitivity whose rationale can be effectively communicated to end-users.

Several templates come to mind when considering temporal dependencies:

1. **Fixed**:
$$class(o,t) = K \ \forall t \geq 0$$
with $K > 0$ constant. The classification remains constant over time.

2. **Step function**:
$$class(o,t) = K_i \ \forall t \in [t_i, t_{i+1})$$
with $0 = t_0 < t_1 < \cdots < t_{n-1} < t_n = \infty$ and $K_i > 0$ constants. The classification changes according to some step function. This includes the case where a previously classified object becomes public knowledge after some specified period of time.

3. **Linear decay**:
$$class(o,t) = \max\{0, K \cdot t + K_0\}$$
with $K < 0$ and $K_0 > 0$ (again K and K_0 constants). This is the simplest case of progressive loss of sensitivity over time.

4. **Exponential decay**:
$$class(o,t) = K \cdot e^{-\alpha t} \ \forall t \geq 0$$
with $K > 0$ and $\alpha > 0$. This is straightforward case of continuous loss of sensitivity over time.

The above are not intended to be exhaustive. Further fundamental templates can be created and combined as desired. For example, sensitivity might be constant for a while and then decay. Sensitivity may also increase over time: the step function can model increasing sensitivity and further fundamental templates can be created to model it continuously.

We now assume that the sensitivity whose temporal trajectory we have just modelled represents some measured and communicable parameter of a distribution. Next we elaborate on how temporal requirements can be integrated with uncertainty modelling.

3.1 Putting the Two Together

We have now defined simple but plausible templates for the temporal evolution of particular sensitivity descriptors and have indicated how at any particular point in time uncertainty in the sensitivity can be modelled using a stretched and offset Beta distribution. We need to put the two together, and this can be achieved in several ways.

In the most general case, the temporal evolution can be specified by a list of time instants and Beta parameters of the form
$$[t_i, (\alpha_i, \beta_i, \gamma_i, \lambda_i)] \tag{6}$$
for $i = 0, 1, \ldots$. The semantics are clear: sensitivity in the interval $[t_i, t_{i+1}]$ is given by a Beta $g(x; \alpha_i, \beta_i, \gamma_i, \lambda_i)$ as defined in expression (3). This allows us to capture in a simple manner any desired variation in time and uncertainty – see, for example, Figure 2(a).

A more compact form can be provided in some cases. The sensitivity distribution over time can be also defined by a Beta of the form
$$g(x; \alpha(t), \beta(t), \gamma(t), \lambda(t)) \tag{7}$$
This allows us to simultaneously model changes in uncertainty and sensitivity. For example, in Figure 2(a) a Beta with parameters $\alpha = \beta = 3$ is initally shifted 5 positions to the right to model a classification "between 5 (SECRET) and 6 (TOP SECRET), with mean 5.5 and a symmetric shape". The template given by functions $\gamma(t)$, $\alpha(t)$ and $\beta(t)$ allows to:

1. reduce exponentially the sensitivity level of the object: the more time elapsed, the less significant the changes; and

2. reduce progressively the amount of uncertainty and approach a delta function.

Figures 2(b) provide another example where the skewness of the distribution is taken into account. This might be useful, for example, to coach requirements of the form "sensitivity should evolve conservatively, i.e. not allowing too much uncertainty on low security levels".

The above merely constitute some illustrative examples. The scheme is sufficiently general as to accommodate many other temporal templates.

4 Integrating Uncertain and Time-varying Sensitivities with Risk-based Access Control

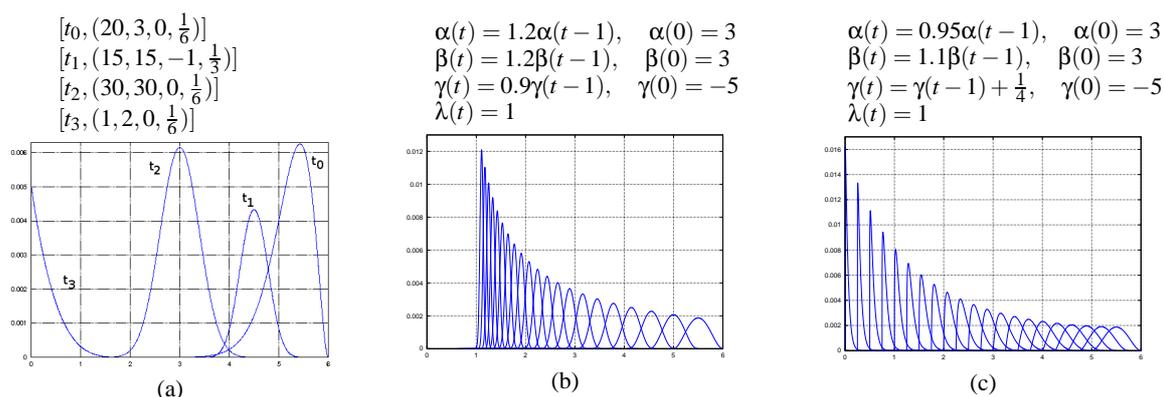We now describe how the templates introduced above can be integrated into an access-control model

$$[t_0, (20, 3, 0, \tfrac{1}{6})]$$
$$[t_1, (15, 15, -1, \tfrac{1}{3})]$$
$$[t_2, (30, 30, 0, \tfrac{1}{6})]$$
$$[t_3, (1, 2, 0, \tfrac{1}{6})]$$

$$\alpha(t) = 1.2\alpha(t-1), \quad \alpha(0) = 3$$
$$\beta(t) = 1.2\beta(t-1), \quad \beta(0) = 3$$
$$\gamma(t) = 0.9\gamma(t-1), \quad \gamma(0) = -5$$
$$\lambda(t) = 1$$

$$\alpha(t) = 0.95\alpha(t-1), \quad \alpha(0) = 3$$
$$\beta(t) = 1.1\beta(t-1), \quad \beta(0) = 3$$
$$\gamma(t) = \gamma(t-1) + \tfrac{1}{4}, \quad \gamma(0) = -5$$
$$\lambda(t) = 1$$

(a) (b) (c)

Figure 2: Examples of templates for time-varying sensitivity with uncertainty. In figures (b) and (c) distributions evolve towards the left, i.e., the rightmost distribution corresponds to the label at time t_0, the next one to t_1, and so on.

based on quantified estimators of risk. Although not covered here, a similar approach could be attempted with trust-based schemes. For the purposes of this paper, we will use the Fuzzy MLS model (Chen et al., 2007a) as a convenient framework. For completeness and readability, we first provide a brief review of the model. Subsequently we show how the proposals above can be integrated within this framework.

4.1 Review of Fuzzy MLS

In (Chen et al., 2007a), $risk$[3] is defined as a function of the "gap" between subject's and object's security level (sl and ol, respectively)

$$risk(sl, ol) = Val(ol) \cdot P(sl, ol) \quad (8)$$

Here $Val(ol)$ is the *estimated value of damage* upon disclosure of the object. The security level is generally considered to represent the order of magnitude of damage, and hence Val is defined as

$$Val(ol) = a^{ol} \quad (9)$$

for some $a > 1$. Note that it is implictly assumed that higher sensitivity corresponds to higher values of the object's security level ol.

The *probability of unathorised disclosure*, $P(sl, ol)$, is defined as a combination of two factors as

$$P(sl, ol) = P_1(sl, ol) + P_2(sl, ol) - P_1(sl, ol)P_2(sl, ol) \quad (10)$$

The first term, $P_1(sl, ol)$, measures the probability that a user with security level sl leaks information of level

[3]This quantifies the risk concerned with the *simple security property* (no read-up) of the Bell-La Padula model. Please see (Chen et al., 2007b) for details about how Fuzzy MLS addresses the concern of the $*$–*property*.

ol by succumbing to *temptation*. It is defined as a sigmoid of the form

$$P_1(sl, ol) = \frac{1}{1 + exp(-k(TI(sl,ol) - mid))} \quad (11)$$

The term $TI(sl, ol)$, called the *temptation index*, roughly indicates how much a subject with security level sl is tempted to leak information with level ol. It is defined as

$$TI(sl, ol) = \frac{a^{-(sl-ol)}}{M - ol} \quad (12)$$

The intuition for the above formulae can be found in (Chen et al., 2007a). The number mid in expression (11) is the value of TI that makes P_1 equal to 0.5, and the term k serves to control the slope of P_1. The value of m is the *ultimate object sensitivity*, and the TI approaches infinity as ol approaches M. (The idea here is that access to an object with sensitivity level M or greater should be granted by a human being and not a machine.)

The second component, $P_2(sl, ol)$, is a measure of the *probability of inadvertent disclosure* for information belonging to a given category, regardless of the object's security level. We shall not elaborate on it, as the extensions proposed in this paper do not affect it directly. Please refer to (Chen et al., 2007a) for further details.

4.2 Integrating with Distributions

Fuzzy MLS assumes that both the subject and object labels are static. We can readily incorporate uncertain and time-dependent sensitivities into the risk estimate as follows. In order to simplify the notation, from now on we will denote by $l_o(x, t)$ the pdf associated with the security level of object o at time t. The variable x

indicates the sensitivity and ranges from 0 to S (e.g., in previous examples we used $S = 6$). The same notation will be used for subjects clearances. Thus the pdf of a user s will be denoted by $l_s(x,t)$

Given a subject s and and object o to be accessed at time t, the tempation index is defined as

$$TI'(s,o,t) = \int_0^S \int_0^S TI(x,y)l_o(x,t)l_s(y,t)\mathrm{d}x\mathrm{d}y$$
$$= \int_0^S \int_0^S \frac{a^{-(x-y)}}{M-y}l_o(x,t)l_s(y,t)\mathrm{d}x\mathrm{d}y$$
(13)

Expression (16) constitutes the natural extension of the TI to a continuous case, where the index is computed over the entire range(s) of sensitivities given by the Beta distributions. Consequently, the probability of unauthorised disclosure, now denoted $P_1'(s,o,t)$, can be computed as in expression (11), although now using $TI'(s,o,t)$ rather than the previous $TI(sl,ol)$.

Regarding the estimated value of damage, expression (9) can be replaced by

$$Val'(o,t) = \int_0^S a^x l_o(x,t)\mathrm{d}x$$
(14)

analogously as it was done for the temptation index.

At a given time t, risk can be computed as before, i.e. weighting the value of damage by the probability of unauthorised disclosure as

$$risk'(s,o,t) = Val'(o,t) \cdot P'(s,o,t)$$
(15)

In a practical implementation, previous expressions can be easily replaced by discrete approximations for convenience.

5 Extending Uncertainty to Contextual Information

It is widely recognised that many access control decisions should depend not only on the identities of the subject and object involved, but also on the context where the access will take place. Thus for example, a user might have unconditional access to a document provided he is at the office and the request is done between 9 am and 5 pm. Context information is often assumed to be publicly available. However, when used as an input to an access control decision it should be properly verified or else ensure (e.g., by cryptographic means) its correctness. *Location*, for example, is usually referred as an important factor when dealing with access control decisions. Ensuring that the location provided by the requester is authentic may not always be an easy affair (see e.g. (Brands and Chaum, 1993; Denning and

MacDoran, 1996; Sastry et al., 2003) for some possible solutions). In some scenarios, measures to guarantee the authenticity of the requester's location may not be available, and therefore some uncertainty will be inevitably present on this information. But uncertainty comes from other sources as well. In a battlefield we may want to associate a security label to each location in a map, in such a way that access to information depends, among other attributes, on the requester's current position. Such labels should not be static assessments of sensitivity, for in a dynamic and unpredictable environment the situation around a position is likely to change over time (e.g. if the enemy moves).

Location only constitutes a particular example of contextual information generally taken into account to grant or deny access to information. In the area of risk-based access control, Cheng and Karger (Chen and Karger, 2008) have identified multiple contextual factors that may contribute to information leakage. These factors consider security-relevant features of the information systems, communication channels, physical environment and human users. In practical terms, specific measures of such factors are interpreted as *risk indices* which, combined together, contribute to assess the global risk.

The templates introduced above to model uncertainty in labels and clearances can be directly applied to context information, particularly in the form of risk factors. If c is a contextual variable (e.g., location, time), a time-varying probability distribution $l_c(x,t)$ can be associated to c. The domain of x is now specific to c (e.g. coordinates in a 2D battlefield, a time interval). We assume that for each c there exists a function $r_c(x)$ mapping each value of x into $[0,1]$, and we interpret this as the *risk* incurred by granting access to a request when $c = x$. Uncertainty in c can now be taken into account as before, so the contextual risk introduced by c is given by

$$r_c'(t) = \int r_c(x)^x l_c(x,t)\mathrm{d}x$$
(16)

Expression (15) should be modified so that contextual factors help to modulate the risk purely derived from the MLS model. We propose a multiplicative scheme of the form

$$risk(s,l,c_1,\ldots,c_k,t) = risk'(s,o,t) \cdot \prod_{i=1}^k r_c'(t) \quad (17)$$

where $c_1,\ldots,c_k$ are contextual variables involved in the decision making.

6 Related Work

The need for access control schemes more flexible than classical approaches has been repeatedly pointed out in recent years, particularly in the context of mobile ad hoc networks. Even though the concept of "risk" is explicitly mentioned by many authors, the great majority of the new models actually rely on a notion of "trust" among parties in order to make access decisions. Trust and risk are indeed related and might be used interchangeably in some contexts, but in an essential sense they are different concepts.

Dimmock *et al* (Dimmock, 2003; Dimmock et al., 2004) explored the relationships between trust, risk and privileges in a trust-based access control setting. Their proposal relies on the idea of granting or denying access depending on the trust it has in the requesting principal and the risk of granting the request. Intuitively, the higher the risk of access, the higher the trust needed in the requester to grant access. In (Dimmock et al., 2004) the authors propose a *quantifiable* definition of risk based on the classical combination of cost and likelihood of outcomes. This model is later discarded in (Dimmock et al., 2004) due, according to the authors, to the "insufficient expressiveness of the risk metrics to capture all the subtleties conveyed by the trust value". Instead, the policy author is provided with a language to express specific rules to compare trust and expected cost information.

Tuptuk and Lupu discuss in (Tuptuk and Lupu, 2007) a very similar idea, namely to use risk to determine the level of trust needed to access a resource. For an authorisation to take place, a measure of trust in the requester needs to exceed a given risk threshold. The risk threshold is acknowledged to be dynamic and mainly dependent on the current context. This work, however, assumes that the metric to obtain such a risk is given.

Diep *et al* propose in (Diep et al., 2007) to make access decisions after a risk assessment of both the request and the context. Risk is estimated for the classical three security properties (confidentiality, integrity and availability), again as a combination of cost and likelihood in a particular context, and then a global risk index is computed.

Though related to the our approach, none of these works explicitly address the notion of risk in an MLS setting.

7 Conclusions and Future Work

Risk-based access control models–and particularly those based on a *quantified* definition of risk,

such as Fuzzy MLS–may be of help to address some of the difficulties that classical schemes are experiencing when dealing with dynamic and unpredictable environments. In this paper we have shown how models such as Fuzzy-MLS can be extended to effectively process uncertain and time-varying security specifications. By explicitly expressing sensitivity as a probability distribution, both security labels and clearances are, in a sense, more accurate in their purpose of reflecting real-world situations. We have also shown how these notions can be extended to contextual information.

In future work we will address questions related to the language needed to express authorisation policies based on risk assessment with uncertainty. Fuzzy logic seems a natural candidate for such a purpose.

In Section 6 we have given account of some recent works exploring the idea of using risk to determine the level of trust required to access a resource. The converse seems not to have been so well studied; namely, can we exploit (quantifiable) trust measures to determine risk? Consider for instance the scenario discussed in Section 2.1, where the (distribution associated with the) label of an object is obtained from the opinions of a number of experts. Such inputs might be somehow weighted by a measure of trust on the subject's organisation, agency, expertise, etc. This and other relationships between access control, trust and risk will be explored in future work.

Acknowledgments

REFERENCES

Bishop, M. (2002). *Computer Security: Art and Science.* Addison-Wesley.

Brands, S. and Chaum, D. (1993). Distance-bounding protocols. In *EUROCRYPT'93*, pages 344–359. Springer-Verlag. LNCS 765.

Chen, P.-C. and Karger, P. (2008). Risk modulating factors in risk-based access control for information in a manet. Technical report.

Chen, P.-C., Rohatgi, P., Keser, C., Karger, P., Wagner, G., and Reninger, A. (2007a). Fuzzy multi–level security: An experiment on quantified risk–adaptive access control. In *IEEE Symposium on Security and Privacy*, pages 222–230. IEEE Press.

Chen, P.-C., Rohatgi, P., Keser, C., Karger, P., Wagner, G., and Reninger, A. (2007b). Fuzzy multi–level security: An experiment on quantified risk–adaptive access control. Technical report.

Denning, D. and MacDoran, P. (1996). Location-based authentication: Grounding cyberspace for better security. *Computer Fraud & Security*, (2):12–16.

Diep, N., Hung, L., Zhung, Y., Lee, S., Lee, Y.-K., and Lee, H. (2007). Enforcing access control using risk assessment. In *Proc. 4th European Conference on Universal Multiservice Networks*, pages 419–424.

Dimmock, N. (2003). How much is 'enough'? risk in trust-based access control. In *IEEE Int. Workshops on Enabling Technologies: Infrastructur for Collaborative Entreprises – Enterprise Security*, pages 281–282.

Dimmock, N., Belokosztolszki, A., Eyers, D., Bacon, J., and Moody, K. (2004). How much is 'enough'? risk in trust-based access control. In *SACMAT'04*, pages 156–162.

MITRE (2004). Horizontal integration: Broader access models for realizing information dominance. Technical report. `http://www.fas.org/irp/agency/dod/jason/classpol.pdf`.

Navy (May 2007). Navy maritime domain awareness concept. Technical report. `http://www.navy.mil/navydata/cno/Navy_Maritime_Domain_Awareness_Concept_FINAL_2007.pdf`.

Sastry, N., Shankar, U., and Wagner, D. (2003). Secure verification of location claims. In *ACM Workshop on Wireless Security*.

Tuptuk, N. and Lupu, E. (2007). Risk based authorisation for mobile ad hoc networks. In *AIMS*, pages 188–191. Springer-Verlag. LNCS 4543.