Australian Government
Department of Defence
Defence Science and
Technology Organisation

# Infoseconomics: A Utility Model for Information Security (U)

*John Yesberg*

Command, Control, Communications and Intelligence Division

Defence Science and Technology Organisation

DSTO–TR–2485

## ABSTRACT (U)

We propose a new model for security based on the utility of information availability. Where certain information should be made available to a certain person, the utility of that access is given a positive value, and where the information should *not* be made available, it is given a negative value. The magnitude of the utility describes the importance of allowing or preventing the access. We describe extensions to the model for time, context, and subjective dependencies, and show how it can be applied in some simple situations.

APPROVED FOR PUBLIC RELEASE

*Published by*

*DSTO Defence Science and Technology Organisation*
*PO Box 1500*
*Edinburgh, South Australia 5111, Australia*

*Telephone: (08) 7389 5555*
*Facsimile: (08) 7389 6567*

*AR No. AR 014–866*
*September 2010*

# Infoseconomics: A Utility Model for Information Security (U)

## Executive Summary (U)

It is very common for computer security policies to express that certain people, or categories of people, should not be able to access certain data. However, in real life, we know that it is much more important to prevent some illegal accesses than other illegal accesses. Existing languages for computer security policies do not give us any way to express this.

A "secure brick" is a system which successfully prevents all illegal accesses to information — it prevents legitimate accesses too, because it does nothing. Most computer security policies would be satisfied by a secure brick, because they don't have language to specify availability requirements.

We propose a new language for specifying computer security policies, which addresses both of these problems. It allows a policy author to specify both (a) how important it is to prohibit some accesses, and also (b) how important it is to allow other accesses. This new language uses a utility model from the field of economics, which prompts the discussion of game theoretic and other economic analyses of secure systems.

Examples are given to show how the utility model might work in considering data communication, networks of different classifications, encryption, and authentiation. Temporal effects are also discussed.

THIS PAGE IS INTENTIONALLY BLANK

# Author

**John Yesberg**
*C3I Division*

John Yesberg has been with DSTO since 1989, and conducted research in computer security since 1990. He was awarded a Ph.D. from the University of Queensland in this field in 1997, and also has degrees in science and engineering, and postgraduate qualifications in management and education. John is a co-author of one of DSTO's "top ten" reports. In 2007 he was awarded a Defence Science Fellowship and was attached to the UK's Defence Science and Technology Laboratory for a year.

THIS PAGE IS INTENTIONALLY BLANK

# Contents

# 1 Introduction

Security requirements are traditionally described by saying which people, groups, or roles should, or should not, have access to particular items or classes of information. A variety of security models from the literature provide different language for expressing these requirements. They may use terms such as discretionary access control, classification and clearance [3], integrity levels [4], constrained data items [8], roles [13], separation of duties[23], or conflict of interest classes [5].

But merely saying whether access should be granted or not is very restrictive. There are many real world situations where we need more expressive language — in particular where some permissions and prohibitions are more important or valuable than others. Existing models do not provide any language for describing these situations.

We note that economists have developed the concept of utility [26] for expressing quantitative preferences. In this paper, we propose a utility-based "Infoseconomic Model" for expressing preferences for different information access permissions — combining economic concepts with those of information security.

The following section presents some of the motivating scenarios which have highlighted for us the need for a quantitative, utility-based model for describing security requirements. Then, we present the core utility-based Infoseconomic Model, and describe some extensions with which we can consider context and temporal dependencies. Next we give some examples of how the model can be used to describe some simple situations. The discussion in Section 5 shows how the Infoseconomic Model relates to a number of other concepts in information security. In the conclusion, we foreshadow some future work.

# 2 Motivations

There are many "real world" security issues which are not expressible in traditional security models. Here we describe some of those that motivated the development of the utility-based model.

## 2.1 Risk Management: Security at Any Price?

In security-critical systems, a system failure may result in a breach of security that might endanger a military mission or national security. Organisations are wise to assess how well they will preserve security, before deciding whether to buy and use such systems. Standards [9] have been adopted to allow the competitive market to develop secure solutions for such purchasers, although this has not been a spectacular success [16].

A problem can emerge, though, if security requirements become absolute and non-negotiable. We do not have "safety at any price" for aircraft, cars, medicines, or school playgrounds — we accept a certain level of risk, trying to make that risk "As Low As Reasonably Practical (ALARP)" [21, 18]. We (as individuals and as a society) appear in general to prefer to allocate our resources toward achieving efficient and fast transport, cheap and rapidly available drugs, and affordable and entertaining playgrounds, even if there is a chance of accidents occurring. Similarly, a "security at any price" policy would be a poor use of resources for accomplishing the mission of national defence. In the same way that a military commander may be required to risk lives of uniformed personnel in conducting a mission, it may be appropriate to consider risking the security of information if, say, adoption of a new technology makes a net positive contribution to achieving the overall mission.

There is much more to say about this argument than is appropriate for this paper, but the key point is that we do not currently have the language with which to express the relative value of security requirements with other requirements. National security policies do provide coarse-grained concepts (such as "Top Secret", "Secret", and "Confidential", and in the United Kingdom "Impact Levels" [25]) that provide limited means for allocating security resources. But we need to be able to compare security requirements with other functional requirements. Systems engineers recognise that a key to successful project delivery is being able to prioritise and negotiate potentially incompatible requirements [20].

## 2.2 Assurance and Strength of Mechanism

Depending on our responsibilities, resources, and anticipated threats, we can implement access control systems with high assurance [9] and high strength mechanisms, or cheaper and lower assurance systems. We can choose whether to use the same password as on the social networking web site, or a different one (that might be harder to remember). We can choose whether to use a short cryptographic key with an older algorithm such as DES, or a longer key with a newer more secure algorithm such as AES. While there are various policies and heuristics that can help us choose for a particular instance, these aren't integrated into any access control models.

## 2.3 Optimistic Security

A number of authors have identified aspects where security policies may be formulated to cater for most, but not all situations [22, 14, 12, 2, 17, 6, 7]. In some unforeseen situations, it may be necessary for people to adjust or bypass security requirements for the overall good of the patient, mission, or organisation.

Individuals may routinely make minor infractions of the rules in some occasions without ever contemplating major infractions. For a real life example, consider a cancer patient who travels some distance to a hospital for chemotherapy, only to find that the consultant doctor hasn't properly signed the prescription. A nurse who is familiar with the patient's treatment regime, rather than calling in the doctor for a signature, or sending the patient home again, simply (and conscientiously) forges the doctor's signature. The same nurse would never forge the signature to, say, acquire recreational drugs for herself.

Policy is written to guide (or prescribe) behaviour within a certain range of foreseeable circumstances. However, occasions that the policy writers did not envisage do sometimes arise. Modern organisational theory (e.g. McGregor's "Theory Y" [19]) suggests that it may be more productive to empower people to make decisions based on guidelines and principles, than to try to encode behaviour in detail in advance. This may lead to a choice to rely less on technical security mechanisms, and more on human ones. If we incorporate people into a socio-technical model of the system, it may be appropriate to use economic mechanisms to enforce security [17, 24, 28, 30].

But to discuss these options meaningfully, we need a model that allows us to compare business, patient, and mission outcomes with the security of information. Simple statements from existing models do not provide this.

# 3 Utility-based Model

We propose a model based on the utility of information availability, called the Infoseconomic Model. This section presents the fundamental *core* of the model: the benefit (or "disbenefit") of a particular piece of information being available to a particular person.

At its simplest, a utility-based or Infoseconomic Model consists of:

- a set of agents $\mathbf{A} = \{A_i\}$,
- a set of information objects $\mathbf{I} = \{I_j\}$, and
- a real valued utility function $U : \mathbf{A} \times \mathbf{I} \to \mathbb{R}$.

For brevity, we may write $u_{ij}$ for $U(A_i, I_j)$.

The value $u_{ij}$ will be positive if there is a requirement that agent $A_i$ should have access to information $I_j$. The greater the value of $u_{ij}$, the higher the benefit (or utility) of $A_i$ having this access. Conversely, if there is a requirement that $A_i$ *not* have access to some information, then the utility $u_{ij}$ will be negative. This expresses that there will be a disbenefit if such access occurs. So the function $U$ is an expression of the security requirements.

When the concept of utility is used in economics, it is usually in the context of looking at the utility of a certain good or service *to a potential buyer*. In the case of the Infoseconomic Model, the utility is from the point of view of whomever might own the information, or be responsible for its security — perhaps the organisation as a whole, such as a commercial company or a military force.

In this paper we have avoided formal definitions of information, agents, and utility. For agents, we have generally imagined a person, but foresee the possibility of referring to automata, roles, and even groups or whole organisations. Similarly, for information, we have been contemplating single information objects as being files. We could equally well consider rows, columns, data items, tables, or entire relational databases, or particular objects (in an object oriented system) to be information objects.

## 3.1 Context Effects

There may be particular events — changes in the overall situation or context — that might lead people or organisations to re-evaluate their desires for information to be available to different agents. The following examples are cases when information availability requirements might change.

- A nation joins a military coalition, and may need access to command and intelligence information.
- A marriage breaks down. One lawyer may no longer act for both parties.
- A student who has boasted in social networks about various aspects of his life may begin to look for work and decide that some previously-public information may no longer be to his advantage.

In cases like these, it may be helpful to model contexts explicitly, in order to describe the effects. We can model the set of possible contexts $\mathbf{C} = \{c_k\}$. The utility would then be a function of the agent, the information, and the context. $U : \mathbf{A} \times \mathbf{I} \times \mathbf{C} \rightarrow \mathbb{R}$. In other cases, there may be no advantage in modelling context explicitly.

## 3.2 Time Dependence

We may, in some cases, prefer to model changes of utility as being merely due to the passage of time. This can relieve us from having to describe individual events. For example, during World War II, the Allies had an exceedingly strong preference not to reveal the breaking of the Enigma cipher to the Germans, keeping it a great secret even from most of the populations of Allied countries. But over time, this preference has changed, and the topic is now discussed openly, even with Germans. While there are certainly particular events that have made this the case, we can abstract these away and assume that it was merely the passage of time that led to the change.

Sometimes, the passage of time is itself the event that leads to changes. For example, in many countries, cabinet minutes and other records of executive governments are published say, 30 years later. Or there may be no need for secrecy about a surprise party, once the party has actually commenced.

Sometimes, we want information to be available continuously. For example, a bank or an Internet business may want to minimise any down-time for its computer systems, lest customers be unable to conduct their business.

In other cases, it is discrete events that have utility. For example, an investor may want a particular “buy” order to be made available to his stockbroker as soon as possible. After the information is transferred and acted upon, there may be no more value in having the information available. This is particularly likely to be the case for an adversary who hopes to exploit the disclosure of sensitive information. It only needs to be communicated once to be of value, and thereafter its utility is negligible.

We have not yet prepared a full formalism to describe the differences between the continuous and discrete event utility. However, a sample application of the latter appears in the next section.

# 4 Sample Applications

## 4.1 Access Control

A security mechanism (or countermeasure) is a component which can be added to a system to provide or improve security properties. For example, a reference monitor[1] is a gateway that only permits certain authorised accesses to proceed. Without the reference monitor, we can imagine that either all access requests would succeed, or perhaps that none would.

If we can work out exactly who has access to different information items in a particular configuration, then we can sum the utilities of the allowed accesses to calculate a net utility of the configuration. This will let us compare different countermeasures, and their combinations.

An idealised mechanism will implement a static policy exactly: it will either provide full (guaranteed) access for agent $A_i$ to information $I_j$ when the policy determines that this is appropriate, and completely prevent this information flow otherwise. Define the policy matrix $\mathbf{P} = \{p_{ij}\}$ by

$$p_{ij} = \begin{cases} 1 & \text{if policy permits access} \\ 0 & \text{otherwise} \end{cases}$$

Then the overall utility of this idealised mechanism (and the associated policy) is the sum[1] of the componentwise products:

$$u = \sum_i \sum_j u_{ij} p_{ij}$$

**Example 1**
*Consider 4 agents $A_1 \ldots A_4$ and a single item of information. Let the utility matrix be:*

$$\mathbf{U} = \left[ \begin{array}{cccc} 3 & 1 & 0 & -4 \end{array} \right]$$

*We can see that the maximum utility would be to have a policy where the first two agents are given access, and the fourth is not. (There is no preference as to whether or not $A_3$ has access in this example.)*

$$\mathbf{P}_{max} = \left[ \begin{array}{cccc} 1 & 1 & 0 & 0 \end{array} \right]$$

*The utility is then 4.*

*A different policy which excludes the second agent, $\mathbf{P} = \left[ \begin{array}{cccc} 1 & 0 & 0 & 0 \end{array} \right]$ will have a lower utility of 3.*

[1] Utility can be defined axiomatically as a total order with an algebra for combining[26], but addition need not be defined. In this respect, it is similar to temperature. In fact, the non-additive nature of utility is very important for illustrating certain phenomena, such as a person being willing to purchase one item (such as a car) for a given price, but not two or more: the marginal utility of the second car would be much less than that of the first car. However, the fact that we can talk about marginal utilities suggests that there is at least some concept of attributing some parts of the total utility to different components. This topic (Aggregation) is discussed further in section 6.

## 4.2 Imperfect Access Control Mechanisms

In some cases, we may need to take account of the fact that no countermeasure is perfect, and there may be various errors in implementation and vulnerabilities to certain threats. It may permit an access where it should not, and it may deny an access that should be permitted. While many security analyses consider the former, it is less common to consider the latter type of error. We moderate the effect of a security mechanism with assurance factors. Instead of using values 0 and 1 as in the policy matrix, we use a probability that the system will provide (or withhold, as appropriate) the information to (or from) the user.[2]

Let the matrix $\mathbb{I}$ be a matrix the same size as $\mathbf{P}$, with all elements 1. If there is a probabilistic assurance of $\rho_{\text{avail}}$ (denoting availability) that the system can provide access to information that is supposed to be accessible, and a probabilistic assurance of $\rho_{\text{conf}}$ (denoting confidentiality) that the system will prevent information flow that is supposed to be prevented. Then the mechanism's matrix $\mathbf{M} = \{m_{ij}\}$ will be constructed from the policy matrix and assurance scalars in this way:

$$\mathbf{M} = \mathbf{P}\rho_{\text{avail}} + (\mathbb{I} - \mathbf{P})(1 - \rho_{\text{conf}})$$

This makes the mechanism matrix $\mathbf{M}$ the same as the policy matrix $\mathbf{P}$ except it replaces the value 1 with $\rho_{\text{avail}}$, and the value 0 with $1 - \rho_{\text{conf}}$. Then the new overall utility is:

$$u = \sum_i \sum_j u_{ij} m_{ij}$$

**Example 2**
*We re-use the utilities from the previous example, and set $\rho_{avail} = 0.9$ and $\rho_{conf} = 0.8$. The mechanism matrix derived from the optimum policy is then*

$$\mathbf{M} = \begin{bmatrix} 0.9 & 0.9 & 0.2 & 0.2 \end{bmatrix}$$

*and the resulting utility is*

$$U = 2.7 + 0.9 + 0 - 0.8 = 2.8$$

*We could compare the hypothetical perfect implementation with $U = 4$ with the real-world one that has $U = 2.8$.*

This allows us to compare the benefits of different individual countermeasures, and with an increase in scale, alternative system configurations.

One of the important aspects of these comparisons is to illuminate decision-making when non-security (functional or other qualities) of the system are affected by different security options. For example, if the choice of a high security system means that certain useful functions would no longer be available, we can calculate the total utility of a system by looking not only at the security utility, but also the functional utilities.

[2] It may sometimes be appropriate for an unsatisfied requirement to have negative, rather than zero, utility.

## 4.3 Encryption

Another mechanism that we might use for security is encryption. In this case, we will need to consider the utilities for different people's access to the ciphertext and the key.

**Example 3**
*Consider a piece of plain (i.e. unencrypted) information $I_p$, which is to be communicated to (or stored for) some internal agents $A_i$, so $u_{ip} > 0$. We also consider some adversary agents $A_a$ who should be prevented from accessing this information, so $u_{ap} < 0$.*

*If we use a symmetric encryption algorithm, then another piece of information to consider $I_k$ is the encryption (and decryption) key. The plaintext $I_j$ is encrypted to become ciphertext $I_c$. Agents $A_i$ need to have access to* both *the ciphertext and the key. If we can allocate most resources to protect the key $I_k$ relatively strongly, but leave the ciphertext $I_c$ more vulnerable to unauthorised disclosure, we can say that $u_{ap} \approx u_{ak} < u_{ac} \leq 0$.*

*If we expect that the adversary may have some chance of success in conducting a ciphertext-only attack against $I_c$, then $u_{ac} < 0$, otherwise we may have $u_{ac} = 0$.*

## 4.4 Information Transfer

In the core of the model, we identified the utilities of information being available to different agents. We now look at ways that agents might communicate information (legitimately or otherwise), and how well such communication may satisfy the requirements.

The simplest example, where information flows from one agent to another, is very similar to the standard access control situation.

**Example 4**
*Consider two agents, $A_1$ and $A_2$, and a single piece of information, $I_1$. We will assume that $A_1$ initially has access to $I_1$ (whether in paper, electronically, or even having heard or deduced it himself).*

*There are two component utilities to consider: $u_{11}$ and $u_{21}$. If $A_1$ keeps the information to himself, then the overall utility will be simply $u_{11}$. If the information is transmitted to $A_2$ either deliberately or accidentally (or if $A_2$ steals it from $A_1$), then the overall utility could be considered (naively) to be $u_{11} + u_{21}$. If $u_{21}$ is positive, then we prefer this latter outcome; if $u_{21}$ is negative, then we prefer the former. (Since $A_1$ has the information in both cases, the utility component $u_{11}$ can be ignored as a constant offset.)*

*If there is a particular probability $\rho$ that $A_2$ will acquire the information, then we can say that the overall expected utility (without the constant offset) is $\rho u_{21}$.*

We can contemplate the extension to a large, possibly infinite, number of agents. A probabilistic method for calculating the likelihood of information flowing from one agent to another appears complex. A simplifying analogy may be to represent the possible communication links as a circuit of nodes, with each inter-agent transmission probability (or rate) $\rho$ corresponding to a resistance $R = \frac{1}{\rho}$. This will allow us to calculate the likelihood that information will be transmitted from one or more agents that already have it to any other agents of interest. Let $\rho_{ij}$ be the probability that agent $A_i$ will

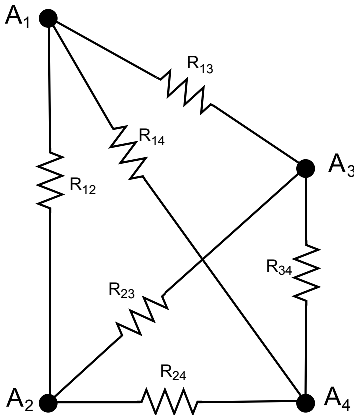

***Figure 1:*** *A circuit analogy for communications.*

transmit information to $A_j$. If $\rho_{ij} = \rho_{ji}$, then a simple resistor network will suffice for all calculations; but otherwise some diodes may be required for asymmetric parts. Graph-theoretic and infinite circuit theory [10] may be useful for further analysis in this area.

**Example 5**
*Consider a network of four agents $A_1 \ldots A_4$. Agent $A_1$ has some sensitive information $I_1$, and there is a negative utility $u_{41} < 0$ if the information is made available to agent $A_4$. By considering the "resistance" each agent may have to communicating information to others, we can calculate the equivalent resistance between nodes $A_1$ and $A_4$ in the circuit in figure 1.*

## 4.5 Short-lived Secrets

Some information is useful and sensitive for short periods of time (perhaps of the order of days), and of little interest afterwards. Examples of this could be military movement orders, and national economic statistics that must be released at precise times to the stock markets. Should such information be treated with the same security as information which must be kept secret for years or decades? One way to model this situation is to assume that there is a certainprobability $\rho$ that information will leak in a given time interval. Using the discrete event utility concept discussed earlier, we are interested in the first time that information becomes available to an adversary.

**Example 6**
*Some sensitive information $I_s$ is being protected from disclosure to an adversary, $A_a$. We may model the likelihood of information leaking on anygiven day (or other time unit) to be constant, $\rho$. Then the probability thatthe information is first disclosed on day $n$ is $(1-\rho)^n\rho$, a geometric distribution [27]. Let $u_{as}(n)$ represent the utility (to the information owner) of the sensitive information being disclosed to the adversary on day $n$. (We expect that $u_{as}(n) < 0$.) Then the* expected *utility is:*

$$E(u_{as}) = \sum_{n=0}^{\infty}(1-\rho)^n\rho \; u_{as}(n)$$

*Rather than using discrete time intervals, we might suppose that security breaches occur at any time and can be modelled as a Poisson process, with a constant average rate of $\lambda$ per unit time. Then the probability density function is an exponential distribution.*

$$E(u_{as}) = \int\limits_{t=0}^{\infty} \lambda e^{-\lambda t}\; u_{as}(t)\; dt$$

*If $u_{as}(n)$ is constant over time $(n)$, then it can be seen for $\rho > 0$ that $E(u_{as}) = u_{as}(0)$. That is, we can expect that the information will eventually be disclosed. However, we might assume that as the information ages, its sensitivity decreases in magnitude from $|u_{as}(0)| = u_{as}^{max}$. Then we can minimise the magnitude of the expected disutility by minimising $\rho$.*

The next example describes an argument for allowing sensitive but perishable[3] information in a domain normally reserved for less sensitive information. Temporal aspects of utility are not traditionally considered when a choice is made about where an item of information should be stored, but the Infoseconomic Model provides language with which we can pose and explore interesting questions.

**Example 7**

*A military organisation has two networks, called Secret (suitable for processing very sensitive material) and Confidential (only suitable for mildly sensitive data). There are stricter controls on the Secret network, and fewer people have access to it, meaning that breaches occur less frequently.*

*Consider some perishable information $I_p$ which we want to be available (at least once) to a recipient $A_r$, but withheld from an adversary, $A_a$: $u_{ap}(t) \leqslant 0 \leqslant u_{rp}(t)$. We need to decide whether to store and transmit the information on the Secret network or on the Confidential network. Traditionally, we would simply look at the maximum disutility $max|u_{ap}(t)|$, and if this is greater than some threshold (see section 5.4), the information would be classified Secret, and hence need to be stored on the Secret network. But given that the information is perishable, maybe we can accept a temporary increased risk if this is outweighed by some other benefit. We will assume that there is a disutility (penalty) $\pi > 0$ for having to access the information on the Secret network. (This penalty may be associated with convenience of access, backup security overheads, inability to transmit to further destinations, and similar inconveniences.) We need to see whether the penalty $\pi$ for using the Secret network is outweighed by the disutility due to the expectation of the information leaking to the adversary, $E(U_a)$. We will model confidentiality breaches on the Secret and Confidential networks as Poisson processes with rates $\lambda_S$ and $\lambda_C$ respectively, and assume that $\lambda_S < \lambda_C$, since breaches are expected less frequently from the Secret network than from the Confidential network. We have chosen to explore adversary utility functions that change linearly from $u_{ap}(0) = u_0 < 0$ at $t = 0$ to zero at a later time $t_0$, $u_{ap}(t_0) = 0$. The expected (dis)utility due to information becoming available to the*

[3] Perishable indicates that the sensitivity of information decays over time.

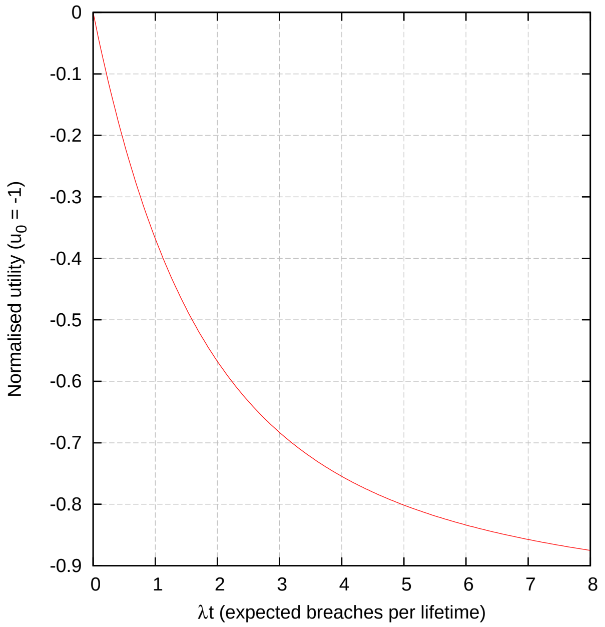

***Figure 2:*** *Expected (dis)utility from perishable information becoming available to an adversary.*

*adversary then becomes (for either network):*

$$\begin{aligned}
E(U_a) &= \int_{t=0}^{\infty} \lambda e^{-\lambda t} u_{ap}(t) dt \\
&= \int_{t=0}^{t_0} \lambda e^{-\lambda t} u_0 (1 - \frac{t}{t_0}) dt \\
&= u_0 \left[ -e^{-\lambda t} - \frac{-e^{-\lambda t}(\lambda t + 1)}{\lambda t_0} \right]_0^{t_0} \\
&= u_0 \big( 1 + \frac{e^{-\lambda t_0} - 1}{\lambda t_0} \big)
\end{aligned}$$

*Notice that* $\lim_{\lambda \to 0} E(U_a) = 0$*. That is, as the number of security breaches per unit time approaches zero, the decreasing likelihood of a breach no longer contributes to any (dis)utility. Similarly,* $\lim_{t_0 \to 0} E(U_a) = 0$*, so the shorter the lifetime of the information, the less we have concerns about it leaking to the adversary. Conversely, as the number of breaches or lifetime of the information increases, we approach the situation where the adversary acquires the information immediately:* $\lim_{\lambda t \to \infty} E(U_a) = u_0$*. Figure 2 shows theexpected (dis)utility as a function of* $\lambda t$*, the expected number ofbreaches during the information's lifetime.*

*Given* $\lambda_S < \lambda_C$*, we will have greater utility (or less disutility) due to breaches if we store the information on the Secret network. But does this increase outweigh the penalty*

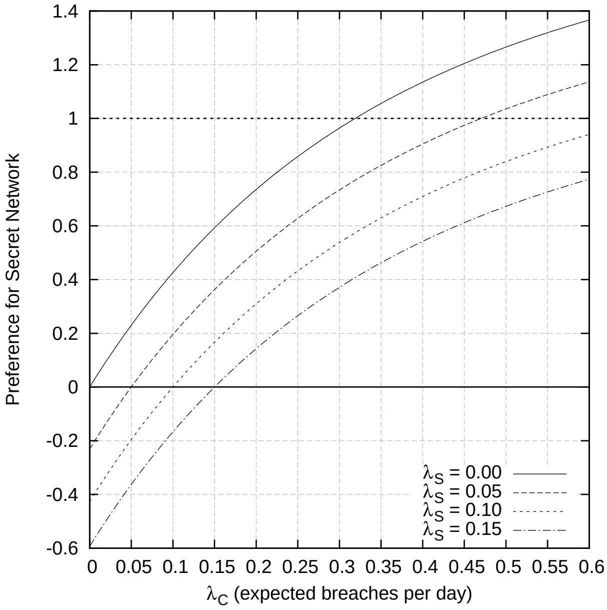

***Figure 3:*** *Breach preference for Secret network as a function of breach rates.*

*$\pi$ associated with that choice? We will choose the Secret network if*

$$\pi < u_0\left(\frac{e^{-\lambda_S t_0} - 1}{\lambda_S t_0} - \frac{e^{-\lambda_C t_0} - 1}{\lambda_C t_0}\right) = P_L$$

*We call the right hand side of this inequality the "breach preference", as it represents the degree to which we would prefer the Secret network because of the expectation of breaches.*

*To look at a numerical example, we will choose the maximum breach disutility to be twice the Secret network penalty, with $u_0 = -2$ and $\pi = 1$, and let $t_0 = 5$ represent the number of days we wish to withhold the information from the adversary. The graph in figure 3 shows the value of the breach preference as a function of $\lambda_C$, for different values of $\lambda_S$. For configurations of $\lambda_S$ and $\lambda_C$ where the breach preferenceis greater than the dashed line $\pi = 1$, the preference to avoid breachesoutweighs the inconvenience, and it could be rational to use the Secret network.*

*On this graph we see that as the $\lambda_C$ increases from left to right, the preference for using the Secret network increases. Also, the preference is higher when $\lambda_S$, the breach rate for the Secret network, is lower. We see also that whenever $\lambda_C < \lambda_S$ the preference to use the Secret network is negative, and whenever $\lambda_C = \lambda_S$ the preference is zero.*

*Whenever the breach preference is above the $\pi$ utility, we prefer to use the Secret network, even considering the inconvenience penalty.*

The assumption that breaches can be modelled as a Poisson allows neatmathematical modelling. A more sophisticated model may look at multiplebreach mechanisms, perhaps each with their own Poisson rates. We can alsoimagine non-Poisson mechanisms, such as latent undetected malware.

## 4.6 Authentication

Authentication mechanisms do not provide 100% confidence that we know exactly who a particular person is. If the outcome of an authentication event is a probability distribution of particular agents, we can quantify the appropriateness of making certain information available based on the various utilities.

For example, if we use a biometric mechanism to identify the current user of a computer, it may provide an indication of which user's template has the best match with the recent measurement. But we may be able to look at the probability that other users (or even unregistered people) could produce such a reading, whether "naturally" or when conducting some kind of attack.

Instead of an access mechanism merely checking whether "the current user" has "permission" to access a given resource, it may be better to look at the expected utility of such an access. Let $\pi_i$ be the probability that agent $A_i$ is the identified user. Then a request to access information $I_j$ could evaluate $\pi_i u_{ij}$ for every user. The mechanism might add all these values to find the net expected utility of the access. Or it might look at the worst case value and compare it with the value for the most likely user. If the most-likely user's utility is outweighed by the worst case risk, access could be denied.

If a system used access and authentication mechanisms in this way, it would be sensible to offer the user an opportunity to re-authenticate, or perhaps a chance to use an additional quantitative authentication mechanism [29] (perhaps an extra password, or a token) to improve the discrimination so that the system is "more confident" of the user's identity.

## 4.7 Modelling the Attacker

We have, so far, always considered utility from a single point of view. However, most economic transactions can only take place when the buyer and the seller of a good have different utilities — the seller would prefer the money to the good, and the buyer would prefer the good over the money. Where we need to consider different people's utilities, we describe the model as being *subjective.*

We need additional notation to describe subjective utilities. We will use a superscript to denote the agent whose viewpoint is being represented, so $u^i_{jk}$ will represent the utility to agent $A_i$ of information $I_k$ being available to agent $A_j$.

At first, we consider a confidentiality attack. Let the owner $A_o$ of some sensitive information $I_s$ have (dis)utility $u^o_{as} < 0$ for the information becoming available to an adversary $A_a$. To the extent that we can measure this utility in monetary units, we can imagine a simple scenario where $A_o$ would be prepared to pay a sum $\kappa < |u^o_{as}|$ to secure the information from such leakage.

In a zero sum game, the disutility to $A_o$ would be equal and opposite to the utility perceived by $A_a$ for the same information. But we need not restrict ourselves thus, and allow $u^a_{as}$ to be different.

To acquire the information, $A_a$ may make an investment $\gamma < u^a_{as}$ in an attack. We will assume there is a range of attack intensities. Attacks with higher intensity are more likely to result in $A_a$ acquiring the information, but they are also more likely to result in $A_o$ or the police noticing the attack and perhaps prosecuting $A_a$.

# 5 Discussion

## 5.1 Model, Policy, and Mechanism

There is not universal agreement about the meaning of, or relationship between, the terms model, policy, and mechanism in computer security. For us, the term model is a way of thinking about a problem; a policy is a codification of how particular agents (whether human or automaton) are to behave in a particular situation; and a mechanism is a process or system for interpreting and executing a policy. A high level policy may be interpreted by a human "mechanism" to generate a lower level policy.

In a large organisation, such as a government department, there may be several layers of policy. The highest level of policy might be a statement about national security, and the lowest level might be a list of which users are permitted to access which files. In our case, the lowest level is likely to be implemented in software, whereas the higher levels are more likely to be implemented by human systems of rules, regulations, procedures, and audits. We may suppose that even the highest level policy (what the "organisation says it wants") may not be exactly what is best for the organisation in every situation. Similarly at the bottom layer, the software implementation may not always execute the policy perfectly — there may be a flaw in the implementation, or it may be attacked by an adversary. The following table is an example of several layers.

| |
|---|
| Accesses that would be good for the organisation. |
| Accesses permitted by the CEO's policy. |
| Accesses permitted by the Project Manager's policy. |
| Accesses permitted by the OS access control list. |
| Accesses actually permitted by the computer. |

Models provide language (concepts and vocabulary) which can be represented in a policy and interpreted by a mechanism. But models can be useful even in situations where none of the policies or mechanisms use these concepts: they can be useful for discussing, for example, why one mechanism might be better than another.

## 5.2 Utility Estimation

We do not imagine or suggest that it will be simple to choose appropriate values for utilities $u_{ij}$. Instead, when trying to populate a model for a real organisation, we foresee that different people will find it difficult to agree on what the values should be.

The problem of estimating utilities can probably be divided into two parts. One part, which we might call the "natural" aspect, involves an understanding of the organisation's processes, and which information is valuable to which people. While achieving the necessary level of understanding may require a great deal of time and effort, we would claim that this is necessary for any thorough security analysis — whether the Infoseconomic Model is to be used or not.

The second "artificial" part of the problem involves choosing a particular numeric scale and zero for the actual numbers to be assigned. This problem is certainly not faced unless the Infoseconomic Model is to be used. At present, we do not have experience modelling anything but trivial situations, and so we are unable to provide real-world-tested guidance to modellers. But we assert that the presence of such a difficulty does not make the model any less useful. In fact, once two analysts understand the model, they can have a *qualitative* discussion about a system or mechanism without having to assign any particular numbers — merely asserting that some utilities are higher than others.

One method that may be useful for estimating utility is to estimate monetary value. In some situations, it may be possible to explore how useful information would be to a particular agent in a particular circumstance by conducting some form of auction. We could find out how much money (or other resource opportunities) a person would be willing to forego to have access to the information. We note that a standard auction would not be appropriate, as this would result in only one person acquiring the information.

Another factor that could be considered is the actual cost of obtaining the information, or in repairing the damage if the information is improperly disclosed. Consider a spy working in a foreign country. If information about the spy is revealed, this might result in deportation or even death, and degradation of an international relationship. To "repair" this damage (in some sense — noting that death is permanent) may take months or years and a great deal of training. Another example might be a military capability (perhaps a weapon, a transport vehicle, or a reconnaissance method) that has a particular vulnerability. If the adversary becomes aware of the vulnerability, the capability may become worthless, and the funds spent will be of no further benefit.

As highlighted in section 3, the value of a secret may decrease over time. In the case of the vulnerability mentioned above, the closer we get to the retirement of that capability, the less it is appropriate to spend on protecting the secrecy of the vulnerability. Loss of the capability one year earlier than planned would cause much less "disutility" than its loss at the beginning of a planned 30 year lifetime.

## 5.3 Confidentiality, Integrity, and Availability

The Infoseconomic Model provides simple language for expression of availability and confidentiality requirements. The concept of their being inverses appears to be novel.

We have not yet discussed integrity requirements — the requirement to prevent unauthorised modification and deletion of data. On one level, which we might call an "information level", we could argue that integrity requirements are really the same as availability. If we require that "the company accounts" are not deleted, and not improperly modified, we could simply say that we require "the company accounts" to be available. Assume that someone improperly modifies a certain spreadsheet or database to introduce errors. Then we could say that "the company accounts" are no longer available, and some new information (perhaps "a misrepresentation of the company accounts") *is* available.

At a “file level”, where we consider files in an operating system (or even a filing cabinet) and read and write operations, then integrity does have a different meaning from availability. Perhaps we should ignore the “information level” and consider only this “file level”. But it is arguably the “information level” at which confidentiality makes the most sense, at least in some cases: it is the information about a trade secret that must be protected, not necessarily a particular file.

It is at this point that we encounter problems with the lack of detailed definitions of “Information Object” used in the model. Consider the case of a person wanting to keep their date of birth confidential from an adversary. The adversary can certainly “know” every possible (or likely) date of birth, but not which date is the person’s birth date. So it appears to be the relationship between the date and the attribute “person’s date of birth” that is the entity that needs protection. If the adversary were to find a file with a single date inside it, he might not know whether that were the person’s date of birth, or some other date. If we notice the adversary with a piece of paper saying “4 April 1960”, we do not accuse him of having stolen confidential information. We could, however, make that accusation if the paper also had “Fred Smith date of birth” on it.

This semantic relationship between the information and its context seems to be related to the difference in meanings of integrity at the different levels. We leave further discussion of this topic to a future paper.

## 5.4 Clearance and Classification

In national security organisations in many countries, it is common to *classify* information with one of a number of labels (such as Confidential, Secret, or Top Secret), according to the (worst possible) degree of damage that could be caused to the national security if the information is disclosed to some unauthorised person.

We can postulate boundary disbenefit levels, $u_{TS} < u_S < u_C \leq 0$. Then the classification $C$ of information $I_j$ could be described as:

$$C(I_j) = \begin{cases} \text{Top Secret,} & \text{iff } \min_i u_{ij} \leqslant u_{TS} \\ \text{Secret,} & \text{iff } u_{TS} < \min_i u_{ij} \leqslant u_S \\ \text{Confidential,} & \text{iff } u_S < \min_i u_{ij} \leqslant u_C \end{cases}$$

Classified information is only permitted to be made available to a person who has the appropriate clearance. We transform this to say that there is a disbenefit for certain classes of agents to have certain information. Let the following subsets represent agents who have particular clearances.

$$\mathbf{C}_{TS} \subset \mathbf{C}_S \subset \mathbf{C}_C \subset \mathbf{A}$$

(The transitivity represents the hierarchical nature of the clearances described so far.)

Then we can say that

$$\begin{aligned}
\min_i u_{ij} \leqslant u_{TS} \quad &\Longrightarrow \quad \forall a_i \notin \mathbf{C}_{TS}, u_{ij} < 0 \\
u_{TS} < \min_i u_{ij} \leqslant u_S \quad &\Longrightarrow \quad \forall a_i \notin \mathbf{C}_S, u_{ij} < 0 \\
u_S < \min_i u_{ij} \leqslant u_C \quad &\Longrightarrow \quad \forall a_i \notin \mathbf{C}_C, u_{ij} < 0
\end{aligned}$$

There is an interesting implication. Consider a document that we wish to make available to some colleagues, but keep confidential from an adversary. Some of the colleagues have Secret clearances, and some have Top Secret clearances. We gradually append more and more sensitive information to the document, and so the utility of its availability to the adversary, already negative, steadily decreases. At some point, the classification of the document transitions from Secret to Top Secret. Before the transition, it is our preference for the Secret-cleared colleagues to see the information. But after the transition, we are required to prevent these people from seeing the information, even though it may be of benefit to the organisation.

This transitional effect is a result of the coarse granularity of the multilevel security system. We could also argue that these colleagues should have had Top Secret clearances. As military organisations increasingly depend on having the right information available to the right people at the right time, they may soon be unable to afford the simplicity offered by multilevel security policies [17].

## 5.5 Aggregation

Standard national security policies (involving classification and clearance) do not cope well with aggregation. This refers to a situation where many pieces of information may individually have a low classification, and yet when they are grouped together, the overall classification of the group is considered to be higher.

In non-military situations, many information-based service providers have to create policies to limit the amount of information that can be accessed. For example, one university library web site says "*Systematic downloading of content is not allowed, including by software such as website crawlers, harvesters or offline browsers.*"

It is tempting to argue that if the disbenefit of agent $A_i$ having access to a set of information items $\mathfrak{I} = \{I_{i_1}, I_{i_2}, \ldots, I_{i_n}\}$ is $u_{ij}$, then the disbenefit of $A_i$ having access to all $n$ items could be simply $\sum_{I_i \in \mathfrak{I}} u_{ij}$. This does seem to provide some reasonable properties for discussing aggregation issues. However, we are compelled to question whether it is meaningful for utilities to be added in this way, given that addition is not part of the formal definition [26] — the law of diminishing returns suggests that the utility of the first item is higher than that of successive identical ones.

If the information items $\mathfrak{I}$ were identical, then presumably the disbenefit of $A_i$ receiving $n$ copies would be no greater than the disbenefit of receiving a single copy[4]. Where

[4] We could imagine that seeing multiple copies might give $A_i$ more confidence in the truth or acceptance of the information. But perhaps this would mean that $A_i$ is looking at the different sources, meaning that the different sources are *part* of the information being considered, and hence each information item is different.

information items are identical or merely overlapping, it seems reasonable to suggest that the value of the set is less than the sum of the values of the individuals.

$$U(A_i, I_{j_1} + I_{j_2}) < U(A_i, I_{j_1}) + U(A_i, I_{j_2})$$

Conversely, it could be that individual pieces of information are worthless on their own, but very valuable in combination. This could occur, for example, if one item was a ciphertext, and the other a key. Where information is *complementary* in this way, we could say that the value of the set is greater than the sum of the value of the individuals.

$$U(A_i, I_{j_1} + I_{j_2}) > U(A_i, I_{j_1}) + U(A_i, I_{j_2})$$

It seems reasonable to imagine that there may be sets of *independent* information items, where the utilities of the individual items can be added together to give the utility of the set.

$$U(A_i, I_{j_1} + I_{j_2}) = U(A_i, I_{j_1}) + U(A_i, I_{j_2})$$

It is, however, difficult to discuss these concepts rigorously without clearer definitions of "information" and "utility". We believe that a more thorough analysis of this situation using an information theoretic background would be appropriate.

## 5.6 Inference

Organisations that gather personal information and then publish statistics are concerned about potential inferences that can lead to breaches of anonymity. They need to compare the utility of smaller statistical granularity (which provides more information to the public) against the increased likelihood that this will allow small groups or even individuals to be identified [15]. Duncan et al. [11] have already discussed the utility of statistical information availability against the potential loss of confidentiality. With the Infoseconomic model, this can be expressed as a utility-utility trade-off, allowing simpler economic decisions.

## 5.7 Economics

Within the field of economics, quantitative utilitarianism and the use of cardinal utility functions seems to have dwindled. It could be asked whether it is sensible to develop a new model for security using old-fashioned economic concepts.

We would argue that the most significant value of a model like the Infoseconomic one proposed here is in the language that it provides, and the additional library of economic concepts it makes available for discussing security problems. It may prove impossible to find objective, accurate numbers for specific information availability utilities. However, this model may allow two people to discuss and describe their requirements and concerns in ways that were not possible without it.

However, there may be more direct economic applications. Governments, commercial, and even non-profit organisations come under increasing pressure from their stakeholders to maximise the effectiveness of their resource allocations. If we can translate the $u_{ij}$ utilities from this model into actual monetary terms, we will be one step closer to determining exactly how much we should spend on security. We explore this aspect in more detail in a forthcoming paper.

The trend in behavioural economics is to assume limited rationality and imperfect information availability. We look forward to exploring these concepts with agent based Infoseconomic Models in the future.

# 6 Conclusion

We have proposed a utility-based Infoseconomic Model as a language for expressing some aspects of real-world security requirements that cannot be adequately expressed by other models. The model assigns a real-valued utility to the availability of information to every possible agent, with a negative utility expressing a confidentiality requirement, and a positive utility expressing an availability requirement. We have shown how it can be used to assess and compare the appropriateness of alternative policies and mechanisms, as well as temporal characteristics.

Several topics have been identified for future exploration.

- A formal treatment of continuous and discrete event utilities.
- A more detailed examination of the precise nature of what it means for information to be confidential, available, and to have integrity. This is likely to involve an understanding of semantic relationships.
- Application of the model to large scale examples, including large networks of communicating individuals, with agent-based and network models.
- An exploration of subjective assessments of utilities. Different agents may perceive utilities differently, even if they try to look from the point of view of the same organisation. This could have implications for the performance of economic security mechanisms.
- The relationship between the desire for confidentiality by one party, the utility of the information to an adversary, and the costs of both countermeasures and attacks. We will speculate that where attacks are targeted against a particular agent, we will need endogenous game-theoretic models, rather than an exogenous threat landscape.

# 7 Acknowledgements

The ideas that have led to this work have come from many conversations over the years. I particularly thank the following for useful discussions: Mark Anderson, Robin Bloomfield, Olivier de Vel, John Fitzgerald, Peter Houghton, Tom McCutcheon, Glennn Moy, Damian Nicholson, Kenny Paterson, Brian Randell, Pankaj Rohatgi,Mark Ryan, Ken Skinner, Abigail Solomon, Colin Thomas.

# References

1. Anderson, J. P. (1972) *Computer security technology planning study*, Technical Report ESD-TR-73-51, HQ Electronic Systems Division (AFSC).

2. Bauer, L., Garriss, S. & Reiter, M. (2008) Detecting and resolving policy misconfigurations in access control systems, *in ACM Symposium on Access Control Models and Technologies*, pp. 185–194.

3. Bell, D. E. & Padula, L. L. (1973) *Secure Computer Systems: A Mathematical Model*, Technical Report ESD-TR-73-278-II, The MITRE Corporation, Bedford, MA.

4. Biba, K. J. (1973) *Integrity Considerations for Secure Computer Systems*, Technical Report MTR–3153, The MITRE Corporation, Bedford, MA.

5. Brewer, D. & Nash, M. (1989) The chinese wall security model, *in IEEE Symposium on Security and Privacy*, pp. 206–214.

6. Cheng, P., Rohatgi, P., Keser, C., Karger, P. A., Wagner, G. M. & Reninger, A. S. (2007) Fuzzy MLS: An experiment on quantified risk–adaptive access control, *in Proceedings of IEEE Symposium on Security and Privacy.*

7. Cheng, P., Rohatgi, P., Keser, C., Karger, P. A., Wagner, G. M. & Reninger, A. S. (2007) *Fuzzy MLS: An Experiment on Quantified RiskAdaptive Access Control*, Technical Report RC24190, IBM.

8. Clark, D. D. & Wilson, D. R. (1987) A comparison of commercial and military computer security policies, *in IEEE Symposium on Security and Privacy*, p. 184.

9. Common Criteria Management Board (2006) Common criteria for information technology security evaluation, version 3.1. Also published as ISO15408.

10. Doyle, P. G. & Snell, J. L. (2006) Random walks and electric networks. http://math.dartmouth.edu/˜doyle/ docs/walks/walks.pdf.

11. Duncan, G. T., Keller-McNulty, S. A. & Stokes, S. L. (2001) *Disclosure Risk vs. Data Utility: The R-U Confidentiality Map*, Technical Report 121, National Institute of Statistical Sciences.

12. Etalle, S. & Winsborough, W. (2007) A posteriori compliance control, *in ACM Symposium on Access Control Models and Technologies*, pp. 11–20.

13. Ferraiolo, D. & Kuhn, D. (1992) Role based access control, *in 15th National Computer Security Conference*, pp. 554–563.

14. Ferreira, A., Cruz-Correia, R., Antunes, L., Farinha, P., Oliveira-Palhares, E., Chadwick, D. & Costa-Pereira, A. (2006) How to break access control in a controlled manner, *in Proceedings of the 19th IEEE International Symposium on Computer-Based Medical Systems*, pp. 847–851. http://www.cs.kent.ac.uk/pubs/2006/2415.

15. Fienberg, S. E. (1994) Conflicts between the needs for access to statistical information and demands for confidentiality, *Journal of Official Statistics* **10**, 115–132.

16. Hearn, J. (2004) Does the common criteria paradigm have a future?, *IEEE Security and Privacy* **2**(1), 64–65.

17. Jason Program Office (1995) *Horizontal Integration: Broader access models for realizing information dominance*, Technical Report JSR–04–132, MITRE.

18. McDermid, J. A. (2001) Software safety: where's the evidence?, *in SCS '01: Proceedings of the Sixth Australian workshop on Safety critical systems and software*, pp. 1–6.

19. McGregor, D. (1960) *The Human Side of Enterprise*, McGraw Hill.

20. Mead, N. R. (2008) Requirements prioritization introduction. https://buildsecurityin.us-cert.gov/daisy/bsi/articles/best-practices/requirements/545-BSI.html.

21. Melchers, R. (2001) On the ALARP approach to risk management, *Reliability Engineering and System Safety* **71**(2), 201–208.

22. Povey, D. (1999) Optimistic security: a new access control paradigm, *in Proceedings of 1999 New Security Paradigms Workshop*, ACM Press, pp. 40–45.

23. Sandhu, R. (1988) Transaction control expressions for separation of duty, *in Fourth Annual Computer Security Applications Conference*, pp. 282–286.

24. Srivatsa, M., Rohatgi, P., Balfe, S. & Paterson, K. (2008) An economic model for securing cross-domain information flows, *in Annual Conference of the Network Information Sciences International Technical Alliance.* http://www.usukita.org/files/Page54.pdf.

25. UK Cabinet Office (2008) Security policy no. 4: Information security and assurance. http://www.cabinetoffice.gov.uk/spf/sp4_isa.aspx.

26. von Neumann, J. & Morgenstern, O. (2004) *Theory of Games and Economic Behaviour*, 60th Anniversary Edition edn, Princeton University Press.

27. Weisstein, E. W. (2009) Geometric distribution, from Mathworld, a Wolfram Web Resource. http://mathworld.wolfram.com/ GeometricDistribution.html.

28. Yesberg, J. (2008) Facilitating rational access control, *in Advances in Computer Security and Forensics.*

29. Yesberg, J. D. & Anderson, M. S. (1995) QuARC: Expressive security mechanisms, *in New Security Paradigms Workshop*, pp. 34–40.

30. Zhao, X. & Johnson, E. (2008) Information governance: Flexibility and control through escalation and incentives, *in Workshop on Economics of Information Security.* http://weis2008.econinfosec.org/program.htm.

Page classification: UNCLASSIFIED

| | |
|---|---|
| **DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION**<br>**DOCUMENT CONTROL DATA** | 1. CAVEAT/PRIVACY MARKING |

| | |
|---|---|
| 2. TITLE<br>Infoseconomics: A Utility Model for Information Security (U) | 3. SECURITY CLASSIFICATION<br>Document (U)<br>Title (U)<br>Abstract (U) |
| 4. AUTHOR<br>John Yesberg | 5. CORPORATE AUTHOR<br>Defence Science and Technology Organisation<br>PO Box 1500<br>Edinburgh, South Australia 5111, Australia |

| | | | |
|---|---|---|---|
| 6a. DSTO NUMBER<br>DSTO–TR–2485 | 6b. AR NUMBER<br>AR 014–866 | 6c. TYPE OF REPORT<br>Technical Report | 7. DOCUMENT DATE<br>September 2010 |

| | | | | |
|---|---|---|---|---|
| 8. FILE NUMBER<br>2009/1041185/1 | 9. TASK NUMBER<br>INT 07/012 | 10. TASK SPONSOR<br>DEPSEC IS&IP | 11. No. OF PAGES<br>20 | 12. No. OF REFS<br>30 |

| | |
|---|---|
| 13. URL OF ELECTRONIC VERSION<br>`http://www.dsto.defence.gov.au/corporate/reports/DSTO-TR-2485.pdf` | 14. RELEASE AUTHORITY<br>Chief, Command, Control, Communications and Intelligence Division |

15. SECONDARY RELEASE STATEMENT OF THIS DOCUMENT

*Approved for Public Release*

OVERSEAS ENQUIRIES OUTSIDE STATED LIMITATIONS SHOULD BE REFERRED THROUGH DOCUMENT EXCHANGE, PO BOX 1500, EDINBURGH, SOUTH AUSTRALIA 5111

16. DELIBERATE ANNOUNCEMENT

No Limitations

17. CITATION IN OTHER DOCUMENTS

No Limitations

18. DSTO RESEARCH LIBRARY THESAURUS

Information Security, Economics, Policies

19. ABSTRACT (U)

We propose a new model for security based on the utility of information availability. Where certain information should be made available to a certain person, the utility of that access is given a positive value, and where the information should *not* be made available, it is given a negative value. The magnitude of the utility describes the importance of allowing or preventing the access. We describe extensions to the model for time, context, and subjective dependencies, and show how it can be applied in some simple situations.

Page classification: UNCLASSIFIED