Journal of Internet Technology, vol. 11, no. 3 , pp. 407-416, May. 2010.

A Probabilistic Approach to Information Control

ÅSMUND AHLMANN NYRE
Department of Computer and Information Science
Norwegian University of Science and Technology
NORWAY
nyre@idi.ntnu.no

MARTIN GILJE JAATUN
Department of Software Engineering, Safety and Security
SINTEF ICT
NORWAY
Martin.G.Jaatun@sintef.no

Abstract

Information aggregation is identified as one of the key threats of the emerging Semantic Web. Although the new web has experienced slow deployment, the advent of Cloud Computing is assumed to boost the rate of deployment by providing means to conduct resource intensive analysis, tracing and querying of semantically annotated and linked data with minimal investment cost for suppliers. However, the threat of information aggregation poses a severe threat to privacy protection, business-critical information and some aspects of national security. The lack of security mechanisms to restrict access to and usage of information after it has been distributed, further contributes to the risk of aggregation since redistribution cannot be controlled. In this paper we propose a probabilistic approach to information control based on trust management systems. Our solution provides the user with a view of the amount of information that any given entity probably has received through redistribution (from others), in order to determine the level of aggregation the entity can perform. We define a middleware architecture and provide an implementation in a simulation environment. Initial results from experiments demonstrate the accuracy of the model and that there may be significant benefits from the approach in several application areas.

Keywords: *Information control, Security, Policy enforcement, Trust-based systems.*

1 Introduction

The Semantic Web as envisioned by Berners-Lee et al. [1] represents a shift from machine readable data towards machine understandable data, allowing for automated intelligent decisions based on the meaning of data on the web. Although we have yet to witness the widespread deployment of semantic technologies, there are several concepts that are gaining quite some momentum, such as the Linked Data Project [1] . The resource intensive processing of such data makes it perfectly suited for the Cloud Computing platform, which is assumed to boost deployment rate for the Semantic Web by significantly reducing the end-user cost of analysis, tracing and querying data [2].

However, with semantically annotated and explicitly linked data, information aggregation may be conducted with far better precision than what is available without these annotations. Although this is generally desired, it also facilitates combining unclassified pieces of information in such a way that the resulting information becomes sensitive. Privacy protection, protection of business information and national security are all application areas in which it is vital to prevent aggregating information.

The problem of protection from aggregation is further complicated by the fact that traditional security mechanisms, most notably access control systems, are commonly unable to restrict information usage outside the originating system. Thus, redistribution of data becomes extremely difficult to monitor and restrict. In this paper we elaborate our probabilistic approach to information management presented earlier [3], where redistribution is inferred based on an underlying trust management system. Our approach allows users to visualize the amount of information individual entities probably have received previously (from redistribution), which serves as input to the decision on whether to distribute more information to that entity.

The remainder of the paper is organised as follows. In Section 2 we give a brief overview of current research and

[1] See http://linkeddata.org

solutions to information control and trust and investigate the current challenges. Next, in Section 3 we outline our approach to information control through integrated trust and information management. We next describe our implementation, experimental set-up and provide key results from simulation runs in Section 4. Our solution is then discussed in Section 5, before we give our concluding remarks and outline further research in Section 6.

2 Related work

Although practical security solutions for the semantic web remain elusive, there is an ample body of relevant security knowledge to draw upon. In the following we provide a brief survey on the state of the art of information control, privacy and trust on the semantic web. Privacy is included here since solutions for controlling personal information may easily be extended to general information control.

2.1 Privacy

All major web sites with user interaction currently provide privacy policies describing how personal information will be handled. The fact that such policies are not understood (or even read) by users, served as one of the main motivations for the early Privacy Enhancing Technologies (PETs) [4-6]. The W3C recommendation Platform for Privacy Preferences (P3P) specification [5] utilises a mark-up language to allow websites to declare their privacy policy in a standardised fashion, which again allow user agents to display the policy in a way users can easily understand. P3P does not provide privacy on its own, but merely helps users make informed decisions about interacting with specific websites. Much of the criticism towards this specification [7] stems from the failure to adhere to privacy standards and regulations.

With semantically annotated policies, agents may conduct privacy negotiations autonomously. Several policy languages have been proposed for both security and privacy policy specification (e.g. [8-10]). By matching users' policies (or preferences) with web services' policies, privacy can be maintained automatically without the need for manual investigation. A review and comparison of current policy languages [11] suggests that policy languages in general are quite expressive, but further work is required especially for improved usage control and minimal information disclosure. Another point being made is the need for user-friendly interfaces and the ability to adapt to changing preferences and requirements.

Unlike other security mechanisms, privacy is also protected by law. Hence, any privacy policy (and preference) should be according to the privacy legislation of the given country. EU directives on privacy protection [12, 13] place requirements on member states' legislation as to how personal information is stored, handled and shared. The P3P specification has been criticised for its lack of support for such legislation. The architecture proposed in [14] uses the principles from the EU directives as a foundation for its legislation compliance. The architecture is capable of mediating between users, websites and legislation, to ensure that all parties' requirements are satisfied. While most privacy enhancing technologies are focused solely on protecting personal information explicitly given by users, this architecture is determined to protect active data (controlled by user, e.g. credentials), semi- active data (partly controlled by user, e.g. sensor data) and passive data (uncontrolled by user, e.g. surveillance cameras).

2.2 Usage control

Park and Sandhu [15] proposed the generic UCON usage control model aimed at being generic enough to encompass both traditional access control, Digital Rights Management (DRM) and trust management. As noted by the authors, privacy management (i.e. controlling personal information) may be seen as the reversed version of DRM, where users are placing restrictions on service providers' use of information. The basic model is built up of subjects, objects, rights, authorisations, obligations and conditions. Subjects and objects are similar to that of other access control mechanisms, with the distinction that their attributes may be mutable, i.e. they may change due to access requests. Rights are not considered static and the existence of a certain right or privilege is determined by a usage decision function upon requesting to invoke it. Authorisations determine whether the subject is allowed to perform the requested operation on the object. Obligations refer to the mandatory requirements a subject must fulfil before or during usage while conditions describe how environmental or system status may influence usage decisions.

2.3 Policy enforcement

Policy enforcement has traditionally (e.g. for access control) been done by a central entity, typically the provider. However, with distributed information and ubiquitous environments, the information provider might be required to enforce restrictions on remote devices. Realising this, Sandhu et al. [16] propose a client-side enforcement strategy based on trusted computing.

The approach taken by Lioudakis et al. [14] is to establish a privacy infrastructure similar to that of public keys (PKI). Service providers implement a Discrete Box functioning as a privacy proxy for end-users. The decision on whether to grant requests for personal information is handled by the containing Policy Decision Point and Policy Enforcement Point (PDP/PEP) of the Discrete Box. Policies considered for such a decision include both statutory, service provider and user policies. The idea is that the service provider's privacy proxy guarantees that all applicable policies (regardless of origin) are met whenever access to personal information is granted. To

prevent misbehaving privacy proxies, the infrastructure is equipped with a set of Privacy Authorities to supervise service providers' adherence to general legislation, user policies and their own specific policies. There are apparent similarities with the Certificate Authority required for the X.509 Certificate infrastructure [17]. Additionally, when applied to the semantic web, each user agent must have its own privacy proxy (Discrete Box), which is a major challenge in terms of scalability.

2.4 Trust management

Trust, and more specifically trust management, has received considerable attention from security researchers over the past years [18], apparently without being able to make a definite impact on services that are actually deployed on the Internet.

The problem with trust is that it takes various meanings in various contexts. In a PKI, a certificate is said to be trusted if the link between the owner entity (e.g. user) and the public key is either known in advance, or is confirmed by a trusted entity. On the current web, the content of a web page is assumed to be trusted if it is provided by a trusted source (as seen by the user). What constitutes a trusted source is not trivially explained.

2.4.1 Definitions

Trust is not easily defined and many definitions exist both within computer science and social sciences [19-21]. Mayer et al. [21] state that organisational studies dealing with trust have been hampered by lack of consensus on contributing factors, trust itself and outcomes of trust. This is supported by a survey of organisational trust [20] and a computer science counterpart [19] where several definitions of trust are listed based on different factors and viewpoints. The common factors of these definitions are vulnerability and risk, implying that the trustor must be vulnerable to the actions of the trustee and that the inherent risk is recognised and accepted in order to call it trust. Mayer et al. argue that it is the recognition of risk that separates trust from confidence, where the latter does not consciously consider the risk involved. Cooperation is another element that may be both a contributing factor and an outcome of trust. Trust may result in cooperation and cooperation may result in trust, but they are not dependent on one another. Entities may be forced to cooperate without any trust relation. Similarly, predictability of entities may be a contributing factor of trust, however only if performance is satisfactory. If always performing badly, predictability may lead to decreased trust [21].

We choose to use the definition from [21], where trust is defined as the willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor and control that other party.

2.4.2 Trust models

As with trust definitions, several different trust models have been proposed over the years, covering different aspects and views of trust. Many of the models that have been proposed have been targeting a very specific use (e.g. e-commerce) and therefore have sacrificed completeness for simplicity, while others have attempted to specify general and somewhat complex trust models.

Mayer et al. [21] focused on a general model. They viewed a trust relation as dependent on the trustor's willingness to trust and the trustworthiness of the trustee (as seen by the trustor). The main factors of trustworthiness were identified as ability, benevolence and integrity. On the trustor's part, disposition to trust and perceived risk were identified as the most influential factors with regards to trust. Furthermore, the outcome of a trust relation (experience) is assumed to influence one or more of the trustworthiness factors and hence the trustworthiness of the trustee.

The work by Marsh [22] was an early attempt to establish a formalism for trust in computer science, and artificial intelligence in particular. The formalism allows agents to compute a trust value based on a set of factors in order to arrive at a trust decision automatically. The complexity of the model makes it difficult to use in practice, however as inspiration the model has contributed greatly to advances in research on trust.

Acknowledging that the complexity of several proposed models does not necessarily give better trust assessments, Conrad et al. [23] proposed a lightweight model for trust propagation. The parameters self confidence, experience, hearsay and prejudice are used to model and assess trust. This computational model also allows agents to compute a trust value to automatically perform trust decisions. The degree of self-confidence deter- mines how much influence own experience and hearsay would have on the computed trust value. The prejudice determines the initial value of experience and hearsay, before experience is accumulated.

In the model proposed by Gil and Artz [24] the idea is to arrive at content trust, where the information itself is used for trust calculation. This allows for a whole new range of parameters (such as bias, criticality, appearance, etc.) to be used when assessing trust in resources. The problem of such parameters is that they require user input, which conflicts with the assumption of agents conducting the assessment autonomously.

2.4.3 Trust propagation

Golbeck and Hendler [25] describe an algorithm for inferring trust and reputation in social networks when entities are not connected directly by a trust relationship. This is done by computing the weighted distance from the source to the sink. Any distrusted entity is not included in the computation since the trust assessments done by such entities are worthless. Guha et al. [26] introduce the

notion of distrust to address the problem of expressing explicit distrust as a contrast to the absence of trust. Absence of trust may come from lack of information to conduct a proper trust assessment, while distrust expresses that a proper assessment have been conducted and that the entity should not be trusted. Furthermore, they argue that distrust could also be propagated and proposes several propagation models in addition to trust transitivity, including co-citation, which is extensively used for web searches.

Huang and Fox [27] claim that not all kinds of trust can be assumed to be transitive. They note that trust based on performance, i.e. an entity performing as expected repeatedly, is not necessarily transitive, while trust based on a belief that the entity will perform as expected often is.

3 Probabilistic policy enforcement

Many of the systems described above assume that entities will always adhere to and enforce their specified policies, either because they are trusted or because there is an infrastructure in place that would not allow them to misbehave (e.g. TPM [16]). As a consequence, enforcement is seen as binary; either it is done or it is not.

While simply assuming that all entities will enforce relevant policies is clearly not a good idea, there are quite some difficulties involved in relying on trusted computing for guarantees.

- Trusted computing requires an infrastructure (hardware and software) for it to work. Hence, any entity that does not comply with this is unable to take part.
- Trusted third parties are needed and are not easily established. Although some have been successfully established for the X.509 Public Key Infrastructure, it is not generally viewed as an unconditional success [28].
- There may be situations where users do want to communicate with entities not part of the trusted infrastructure, even though this would generally conflict with their security requirements. Users would therefore be forced to disable the enforcement system in order to communicate.
- With any such system, there is a critical mass of users/providers that must be attained before users will view investments in such tools beneficial.

One of the main problems of relying on trusted infrastructures to restrict information is the fact that users' security policies are highly dynamic and greatly dependent on context. Hence, there may indeed be situations where users would want to interact with entities that normally do not fit with their security requirements (and that are not part of the trusted infrastructure). Users could therefore greatly benefit from an information control system that would:

- Alert the user of the trustworthiness of the other entity.
- Record the user's willingness to interact and the willingness to (potentially) be vulnerable to exploit.
- Provide means to mitigate the risk and calculate criticality and consequence of interaction (e.g. redistribution of data).

We therefore propose a probabilistic approach to policy enforcement, where users are given a probability that their requirements will be respected and polices enforced. Thus when interacting with entities who are known to be less trustworthy, policy adherence is given by a probability metric that the entity will actually enforce it. Our model is concerned with how to handle uncertainty in enforcement and provide a tool for interacting with non-conforming entities while minimising the risks involved. Our intention is to complement trusted infrastructures-based and authority-based information control, rather than replace them.

3.1 Data Recorder

Information aggregation is one of the prominent new threats on the semantic web. With improved ability to extract and combine information it is vital to keep track of what and to whom information has been distributed. In our model, this is done by the Data Recorder (DR), which basically records what data is transmitted to which receivers. This will allow users to refrain from providing more information if what is already known to the receiver can violate the security policy.

The Data Recorder is concerned with recording information pieces, denoted documents in our model. These documents are assumed to have two main properties: size and sensitivity. Size refers to the amount of information (e.g., number of pages, words, bytes), while sensitivity is used to denote the degree to which this information must be protected (e.g., classified, secret, top secret). Based on these properties we can calculate the information value of a document *d* as:

$$V(d) = \text{size}(d) * \text{sensitivity}(d) \qquad (1)$$

Here, the value of a document is used to indicate the relative protection level required and the relative attractiveness as seen by an adversary. Small documents of unclassified information are thus assumed to have relatively low value, whereas large documents of top-secret information are assumed to have relatively high value. Our model does not specify how to compute size and sensitivity of documents, nor does it specify the scale to use for these purposes. This is deliberate, since these properties may vary considerably in different situations.

The DR allows the user to view himself through the eyes of the receiving party, and thereby perform aggregation to see whether too much information is provided.

3.2 Inference Engine

A major problem with distributing information is the fact that control of information is transferred along with the document. But, rather than attempt to control the receiver's use of the information our model focuses on the ability to predict the receivers actions. This is the responsibility of the Inference Engine (IE), which thus constitutes the core of our probabilistic information control model. Although the inference engine could be used to compute probability of any action taken by the receiving entity, our focus in this paper is on redistribution of information.

To infer redistribution of documents, the inference engine must be able to compute the probability that the initial receiver forwards the document and next to whom the document is forwarded. It is important to notice that such forwarding is not necessarily considered illegal or immoral. In a business setting it is quite common for business propositions sent to a specific employee of a company to be forwarded internally, such as to colleagues, superiors, etc. Therefore, one cannot generally infer that trustworthy users never will redistribute documents.

Still, we find it natural that the level of trust the sender has in the receiver affects the probability of redistribution. Increasing level of trust yields decreasing probability of redistribution. We define the redistribution probability $P_r(r)$ of the receiver *r* to be:

$$P_r(r) = 1 - T(r) \qquad (2)$$

Where $T(r)$ denotes the trust the sender has placed in the receiver and is a real value in the range [0, 1]. As can be seen, when the trust level tends to 1, i.e. complete trust, the redistribution probability tends to zero. Although this may be seen as contradicting our initial statement that redistribution may occur despite the fact that the receiver is trusted, we argue that complete trust (i.e. T (r) = 1) would be an extremely rare case. For all practical purposes the probability will be non-zero, i.e., $P_r(r) > 0, \forall r \in R$, where R denotes the set of potential receivers. This is further debated in Section 5.

Predicting the receiver of redistributed documents is perhaps more challenging. We cannot assume that the originator of the document will be able to identify all potential individuals the initial receiver might forward the document to. And, even if we could, it would not scale very well and would be extremely difficult to maintain. We therefore need to view the possible receivers on a higher level. We propose to use groups as a way to identify individuals that might share (i.e. redistribute) information with each other. A group could be formed based on interest (e.g. "conspiracy believers"), on affiliation or demographic data (e.g. sex, age). There are of course numerous ways of forming groups and the receiver will not know all of them. The benefit of predicting redistribution to groups rather than individuals is that group membership may be used to infer previous knowledge when interacting with a previously unknown entity. However, problems arise when users switch groups (e.g. affiliation), since we would need to know the time period the user was member of the group in order to accurately infer what documents probably have been received. Although this is a valid remark, we have aimed for simplicity in our model and chosen to only regard current group membership when determining the potential receivers of redistributed documents.

Information that is made publicly available on the Internet would generally be considered to be available to all. This is handled in our model through the use of a base group where all entities are members. Any information available to the general public will be added to this group.

3.3 Trust Assessment Engine and Trust Monitor

From Section 2.4 we acknowledge that numerous definitions of trust exist. However, in our model the term trust is used to denote the degree to which the trustor believes that the trustee will manage the information according to the policy given by the trustor. Hence, integrity and benevolence, rather than ability, is considered the main factors of trustworthiness.

The Trust Assessment Engine (TAE) is responsible for calculating trust values of different entities in order to determine their trustworthiness. The TAE is thus focused solely on assessing communicating parties and does not take into account risk willingness, vulnerability and criticality.

The trust monitor (TM) is responsible for detecting events that might affect the perceived trustworthiness and the willingness to take risks. The trust monitor is thus responsible for calculating and deciding on what is an acceptable trust level, given the circumstances. Any computed trust value and feedback received from cooperating entities is stored in the trust assessment repository.

We do not propose any specific trust model to be used for in our model, nor do we assume any property of trust models other than the basic features of assessing and monitoring trust. It is worth noting that trust models used in our model should be consistent with our use of the term, such as the model used for simulation (see Section 4).

3.4 Policy decision point

The Policy Decision Point (PDP) is, as the name implies, where decisions are made regarding whether to send information to the receiver or not. The decision is based on input from the other components in the system as well as the overall security policy of the sender. In our case,

Parameter	*Value*
Document rate	
Total information threshold ($I_{\max}$)	100
Document information threshold ($V_{\max}$)	20
Redistribution threshold ($E_{\max}$)	10
Document sensitivity range	1-5
Document size range	1-5
Number of persons	10
Simulation period	1000
Sampling rate	1/10

a) Information control settings

Parameter	*Value*
Initial trust	1.0
Prejudice	0.7
Performance	0.9
Self confidence	0.4

b) Trust settings

Table 1: Simulation settings

considering redistribution of information, the decision will ultimately be based on what documents have previously been sent to the recipient, what information might have been forwarded to him previously, the probability that he would redistribute the document and of course the trust the sender places in him. There are numerous policy languages available such as WS-SecurityPolicy [29], Ponder [9], Rei [30], etc. See [11] for a survey on policy languages and their capabilities. Similar to the trust model mentioned earlier, our model makes no assumptions on how policies are specified such that any of the available policy languages may be used. Since details of policy languages are outside the scope of this paper we have chosen to use threshold values as a way to express redistribution policy.

We define three main points that must be evaluated by the PDP in order to assess whether the redistribution policy is met. In the following we let *s* denote the sender, *r* the receiver and *d* the document to be sent. Firstly, the information value of a document must not exceed the trust-weighted predefined document information value threshold $V_{\max}$. That is,

$$V(d) \leq V_{\max} \cdot T_s(r) \qquad (3)$$

Where $V(d)$ is the information value of a document *d* and $T_s(r)$ denotes the trust the sender *s* has placed in the receiver *r*. The effect of this requirement is that reduced level of trust reduces the maximum allowed information value correspondingly. Secondly, the information value of all previously distributed documents to the receiver, including documents that are assumed to have been redistributed by others, must not exceed the trust-weighted predefined threshold $I_{\max}$. That is,

$$\sum_{d_i \in A_s(r)} V(d_i) + \sum_{d_i \in B_s(r)} V(d_j) \leq I_{\max} \cdot T_s(r) \qquad (4)$$

Where $A_s(r)$ denotes all documents that have been sent to the receiver *r* and $B_s(r)$ denotes all documents that probably have been redistributed by others to the receiver *r*. The effect of this requirement is identical to that of (3) in that reduced trust yields a correspondingly reduced maximum information threshold. Thirdly, and finally, the redistribution probability-weighted information value of the document must not exceed the predefined redistribution threshold $E_{\max}$. That is,

$$P_r(r) \cdot V(d) \leq E_{\max} \qquad (5)$$

Where $P_r(r)$ is the probability that *r* will redistribute the document (see (1)). This requirement ensures that the information given to receivers that are likely to redistribute information will be greatly reduced.

4 Implementation and experimentation

In this section we will describe the implementation of our model for the purpose of simulation and general experimentation. The implementation does not provide any real middleware that can be utilized by other applications, but rather provide a means to see how the model conceptually could work. The implementation is done using the Python programming language and the SimPy discrete event simulation environment [2]. The source code and detailed simulation settings are available from the authors on request.

4.1 Assumptions and simulation set-up

The model we proposed in Section 3 is quite general, allowing for adaptation and configuration to suit different needs. In our implementation we have made some assumptions to be able to implement and experiment with our model. In the following we identify these assumptions and explain briefly their impact on the simulation to be conducted.

- The lightweight model proposed by Conrad et al. [23] was selected as the underlying trust model, with the same configuration as in their experimental set-up.
- Events required to produce experience and subsequently to assess trust relations was simulated as generated interactions between randomly chosen entities. The outcome[3] of these interactions was

[2] Available from http://simpy.sourceforge.net
[3] Called "immediate experience" in [23]

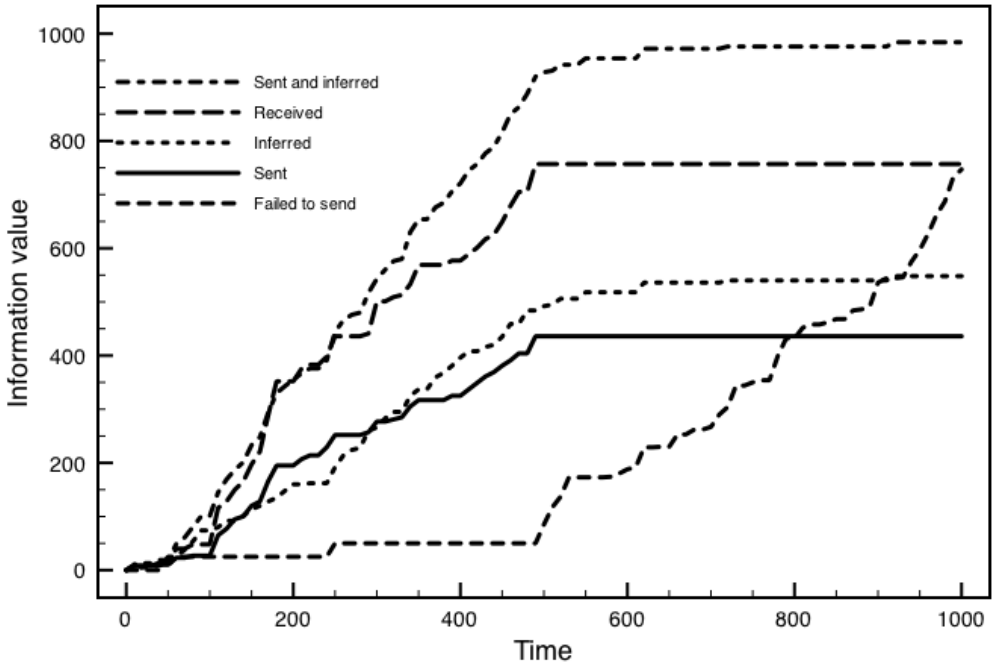

Figure 1: Information distribution between a randomly selected pair of entities

assumed to be binary, either success (= 1) or failure (= 0).

- The evaluation of performance is independent, such that participants interacting do not affect each other's evaluation, and the success rate is specified before hand. Performance is assumed to follow a normal distribution with a predefined mean and variance (see Table 1).
- Documents to be sent are assumed to be generated with random size and sensitivity at a predefined frequency.
- Statistical data from simulations are recorded at a predefined sampling frequency.
- Redistribution is assumed to be done upon reception of a document, and not at a later point.

Simulation parameters were set according to Table 1. Ten persons were set to generate and attempt to send documents to randomly chosen recipients at a rate of one document every tenth time unit. Document size and sensitivity are chosen uniformly from their respective ranges. Statistical data is collected at a frequency of one every tenth time unit. Included in this data collection is the total information value of all documents that have been sent, received and redistributed. In addition, also the documents that failed to be sent and the documents that probably have been redistributed are included. The data is grouped according to sender and receiver. The data forms the basis for our analysis in the upcoming section.

4.2 Metrics

The purpose of the simulation is to show how our model may be used to predict information distribution in an environment where there are no dependable enforcement strategies. We identify one metric for evaluating the effectiveness of our solution and one to show the relative performance compared to models without inference.

In the following we let *s* and *r* be sender and receiver, respectively. Further we let $V_A(s,r)$ denote the total information value transferred directly from *s* to *r*, and let $V_B(s,r)$ denote the additional information value that *s* assumes received by *r* through redistribution. Finally, we let $V_T(s,r)$ denote the total information value originating from *s* that is actually received by *r*. We then define the inference accuracy metric to be $M_{\text{INF}}(s,r)$:

$$M_{\text{INF}}(r,s) = \frac{V_T(r,s)}{V_A(r,s) + V_B(r,s)} \qquad (6)$$

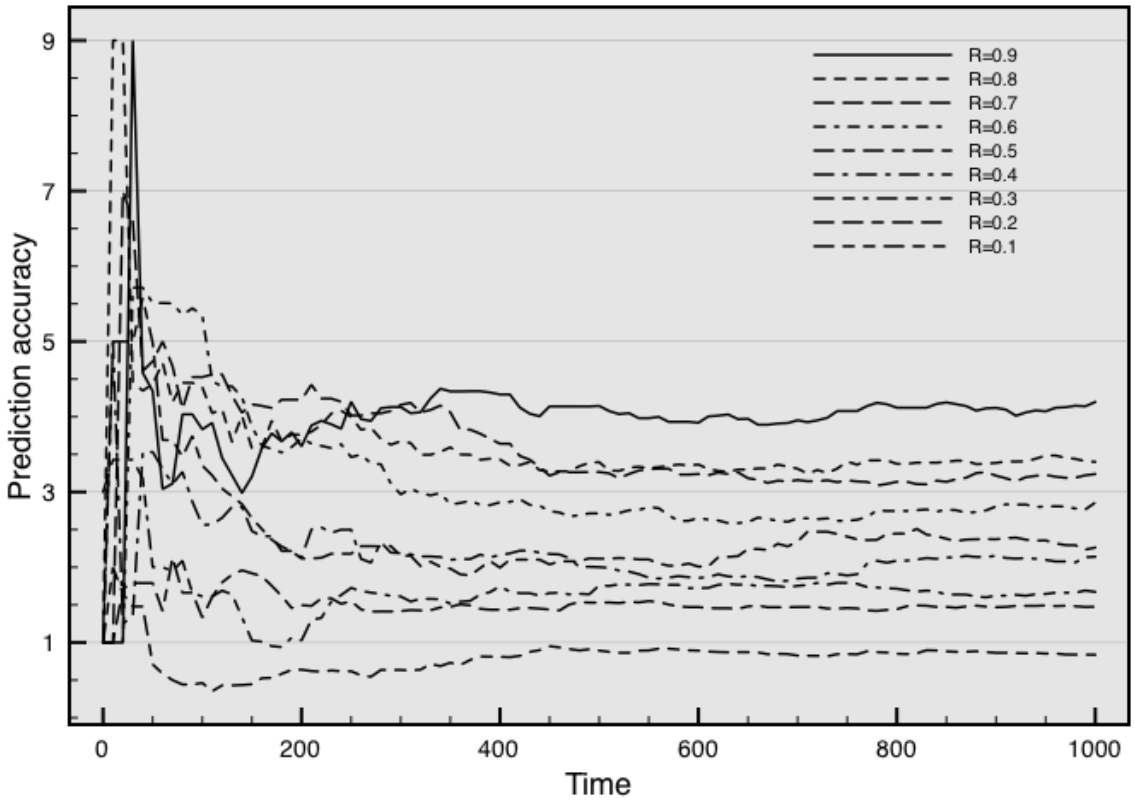

Figure 2: The prediction accuracy metric for varying redistribution rates

We know that the ideal amount of inferred information is given by $V_B(r,s) = V_T(r,s) - V_A(r,s)$, since by inference we want to get as close to the amount of actually received information as possible. This gives us the perfect accuracy at $M_{\text{INF}}(r,s) = 1$. To form a baseline to which we can compare the accuracy metric defined above, we also define a metric for the non-inference accuracy. That is, the degree to which our view on distributed information is accurate if we do not use any inference. This is identical to a system were all distributed information are recorded, and no redistribution is assumed to take place. The metric is identical to that of Equation 6 when setting $V_B(r,s) = 0$, which gives us:

$$M_{\text{NO-INF}}(r,s) = \frac{V_T(r,s)}{V_A(r,s)} \qquad (6)$$

The non-inference metric is defined by the ratio of received information to the directly sent information.

4.3 Results and evaluation

The simulation was run with the settings described above. Figure 2 shows the key parameters that were recorded during simulation for a randomly selected pair of entities. We can see that the as the sum of sent and inferred information tends to the trust-weighted threshold, the sending rate is decreasing and the failed to send-rate is increasing. This is expected since the sending entity will stop once the sum of sent and inferred information exceeds the weighted threshold. The small earlier increases in this rate are due to the fact that sending will fail if the document value exceeds the weighted threshold. We can also see that the sent and inferred information is very closely related to the actual received information until 400 time units. From then on, the difference is substantial.

To get deeper into the accuracy metric, simulation was repeated with varying redistribution rates. The total information threshold was set to positive infinity, to prevent low trust scores (and consequently low thresholds) form reducing the sampling space substantially. Other parameters were set as described in

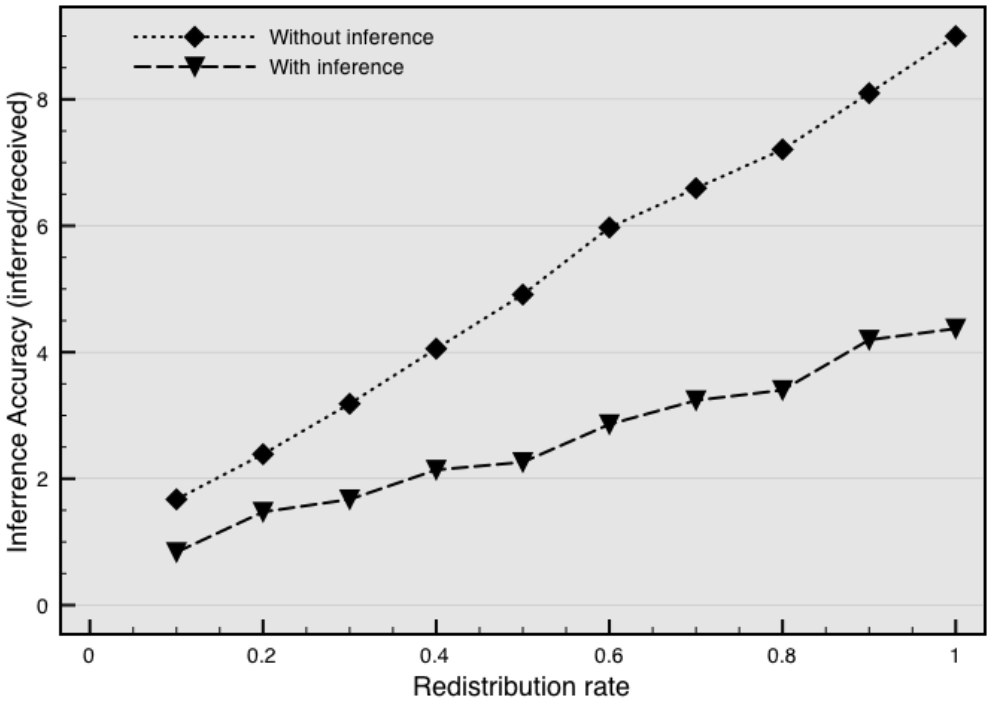

Figure 3: A comparison of the accuracy metric for prediction and non-prediction, for varying redistribution rates

Section 4.1. At each sampling point, the inference accuracy M_{INF} (see (6)) was computed. Figure 3 shows how the inference accuracy varies over time for the different redistribution rates. Not surprisingly, the greater the redistribution rate, the more the accuracy metric deviates from the perfect value (= 1). Since the trust model is set with an expected performance rate of 0.9, we are expecting a redistribution rate in the range of [0, 0.2]. It is only for redistribution rate of 0.1 that the accuracy metric drops below 1, i.e. that the model predicts more information has been sent then what actually is the case. In a worst-case scenario (redistribution rate is 0.9), four times as much information as predicted is actually received.

To evaluate whether the accuracy of our simulated model actually constitute an improvement, the accuracy metric of the inference approach was compared to the accuracy metric without inference. Figure 4 shows the two metrics for varying redistribution rates. It can be seen that the accuracy metric of the inference approach is about half of that without for all redistribution rates. It may be argued that this will always be the case whenever the redistribution rate is above zero, since some inferred information will always be better than none.

5 Discussion

The practical application of information control systems is limited by the human cognitive capacity – or rather, the lack thereof. However, even on the semantic web, information is often ultimately communicated from one human to another, and thus if we want to apply trust to this equation, we have to base ourselves on human notions of trust, which are neither binary nor straightforward.

In fact, the word "trust" is used to mean many things even in a human context, and is often misunderstood when applied to end-user applications. One example is the Pretty Good Privacy (PGP) program, which allows users to collect public keys of friends and associates, and subsequently assign a "trust level" to these keys. A common misconception is that this trust level reflects on the degree of certainty that the given key is the correct key for that particular associate; while in reality it reflects to which degree the user is willing to trust other keys signed by that particular associate. These concepts are quite different: While I am confident that I have the right key for my friend Bob on my computer, I also know that Bob is a fool who at regular intervals is sending money to Nigerian princesses, and also freely signs any key that comes his way.

Our model relies heavily on the idea that there is a negative correlation between level of trust and security

policy compliance. That is, the more a sender trusts the receiver, the less likely the receiver is to violate the security policy[4]. It therefore becomes utterly important to select an underlying trust model that properly reflects this, to prevent misconceptions as described above.

Complete protection from information aggregation requires control of passive and semi-active data. That is, data that is collected by third parties either with or without the subject's cooperation, e.g., interview, observation, interpretations of information. It is apparent that providing control of every possible information source regarding a subject is quite difficult, not to say impossible. Our approach does not attempt to handle secondary sources of information, only information originating from the subject is considered when inferring the redistribution of information. However, should such information from secondary sources be made available to the Data Recorder (e.g., through data mining), then the model will be capable of utilising this when assessing total information distribution.

The assumption is that ultimately everyone[5] violates security policies (or redistributes information); it is merely a question of how much or at what rate. If the assumption is proven invalid there will be no policy violations and consequently no need for inference. However, we argue that although there may be situations were this assumption does not hold, there will certainly be situations where it does hold. Thus, whether the solution proposed in this paper is a good one, greatly depends on the scenario and intended usage.

6 Conclusion and further work

In this paper we have outlined existing approaches to information control and trust management and the fundamental challenge of the emerging semantic web. We have proposed a new way of handling distributed information remotely, based on computing the probability that the recipient will adhere to the established policies. The probability is computed on the basis of trust assertions, user's willingness to trust, and the information involved. We believe that such an approach would facilitate a gradual deployment of software since it may prove beneficial to users, regardless of whether other users have adopted it.

Our simulation shows that the approach could be quite accurate provided that the trust management system correctly captures trust in policy adherence, and of course that the key assumptions, as discussed above, is met. However, we do acknowledge that empirical research is required to determine how the inference engine and policy decision point can be improved to better mimic human behaviour. Also, effort must be placed in how to generalize the implementation to support a wide range of security policy issues (not just redistribution) and how to describe the correlation of trust and policy adherence in general.

[4] Or to redistribute information, as in the simulation scenario

[5] Except those that are fully trusted, i.e. trust= 1

References

[1] T. Berners-Lee, *et al.*, "The Semantic Web," *Scientific America*, pp. 34-43, May 2001.

[2] P. Mika and G. Tummarello, "Web Semantics in the Clouds," *IEEE Intelligent Systems*, vol. 23, pp. 82--87, 2008.

[3] Å. A. Nyre and M. G. Jaatun, "Privacy in a semantic cloud: What's trust got to do with it?," in *Proceedings of the first International Conference on Cloud Computing (CloudCom)*, Bejing, China, 2009, pp. 107-118.

[4] H. Burkert, "Privacy-enhancing technologies: typology, critique, vision," in *Technology and Privacy: The New Landscape*, P. Agre and M. Rotenberg, Eds., ed. Cambridge, MA, USA: MIT Press, 1997, pp. 125--142.

[5] L. Cranor, *et al.*, "The Platform for Privacy Preferences 1.0 (P3P1.0) Specification," ed: W3C Recommendation, 2002.

[6] I. Goldberg, *et al.*, "Privacy-enhancing Technologies for the Internet," presented at the Proc. of 42nd IEEE Spring COMPCON, 1997.

[7] EPIC, "Pretty Poor Privacy: An Assessment of P3P and Internet Privacy," Electronic Privacy Information Center, June 2000.

[8] P. Bonatti and D. Olmedilla, "Driving and monitoring provisional trust negotiation with metapolicies," presented at the Policies for Distributed Systems and Networks, 2005. Sixth IEEE International Workshop on, 2005.

[9] N. Damianou, *et al.*, "The Ponder Policy Specification Language," in *Policies for Distributed Systems and Networks*. vol. 1995, ed Berlin: Springer Verlag, 2001, pp. 18--38.

[10] L. Kagal, *et al.*, "A policy based approach to security for the Semantic Web," in *Semantic Web - ISWC 2003* vol. 2870, ed. Berlin: Springer-Verlag Berlin, 2003, pp. 402-418.

[11] C. Duma, *et al.*, "Privacy in the Semantic Web: What Policy Languages Have to Offer," presented at the Policies for Distributed Systems and Networks, 2007. POLICY '07. Eighth IEEE International Workshop on, 2007.

[12] EU, "Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of such Data," ed: Official Journal of the European Communities, 1995.

[13] EU, "Directive 2002/58/EC of the European Parliament and of the Council Concerning the Processing of Personal Data and the Protection of Privacy in the Electronic Communications Sector," ed: Official Journal of the European Communities, 2002.
[14] G. V. Lioudakis, *et al.*, "A middleware architecture for privacy protection," *Computer Networks*, vol. 51, pp. 4679--4696, November 2007.
[15] J. Park and R. Sandhu, "The UCON-ABC usage control model," *ACM Transactions on Information Systems Secuity*, vol. 7, pp. 128--174, 2004.
[16] R. Sandhu, *et al.*, "Client-side access control enforcement using trusted computing and PEI models," *Journal of High Speed Networks*, vol. 15, pp. 229--245, 2006.
[17] R. Housley, *et al.*, *RFC 3280: Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile*: RFC Editor, 2002.
[18] V. Varadharajan, "A Note on Trust-Enhanced Security," *Security \& Privacy, IEEE*, vol. 7, pp. 57--59, 2009.
[19] D. Artz and Y. Gil, "A survey of trust in computer science and the Semantic Web," *Web Semantics: Science, Services and Agents on the World Wide Web*, vol. 5, pp. 58--71, June 2007.
[20] G. Bigley and J. Pearce, "Straining for Shared Meaning in Organization Science: Problems of Trust and Distrust," *Academy of Management Review*, vol. 23, pp. 405--421, 1998.
[21] R. Mayer, *et al.*, "An Integrative Model of Organizational Trust," *Academy of Management Review*, vol. 2, pp. 709--734, 1995.
[22] S. P. Marsh, "Formalizing Trust as a Computational Concept," Department of Computing Science and Mathematics, University of Sterling, 1994.
[23] M. Conrad, *et al.*, "A lightweight model of trust propagation in a multi-client network environment: to what extent does experience matter?" presented at the Availability, Reliability and Security, 2006. ARES 2006. The First International Conference on, 2006.
[24] Y. Gil and D. Artz, "Towards content trust of web resources," presented at the Proceedings of the 15th international conference on World Wide Web, New York, NY, USA, 2006.
[25] J. Golbeck and J. Hendler, "Accuracy of Metrics for Inferring Trust and Reputation in Semantic Web-based Social Networks," presented at the 14th International Conference on Knowledge Engineering and Knowledge Management, Northamptonshire, UK, 2004.
[26] R. Guha, *et al.*, "Propagation of trust and distrust," presented at the WWW '04: Proceedings of the 13th international conference on World Wide Web, New York, NY, USA, 2004.
[27] J. Huang and M. S. Fox, "An ontology of trust: formal semantics and transitivity," ed: ACM, 2006.
[28] J. Lopez, *et al.*, "Why have public key infrastructures failed so far?," *Internet Research*, vol. 15, pp. 544-556, 2005.
[29] A. Nadalin, et al., "Web Services Security Policy Language (WS-SecurityPolicy)," OASIS2007.
[30] L. Kagal, et al., "Authorization and privacy for semantic Web services," Intelligent Systems, vol. 19, pp. 50-56, 2004.

Åsmund Ahlmann Nyre received his MSc degree in Communication Technology from the Norwegian University of Science and Technology (NTNU) in 2005. He worked until 2007 as Security Engineer at the Norwegian road-tolling company Q-Free, after which he joined SINTEF ICT. As of July 2009, Mr. Nyre holds additionally a position as PhD candidate at NTNU. His research interest includes information control, privacy and network security.

Martin Gilje Jaatun received his MSc degree in Telematics from the Norwegian Institute of Technology in 1992. He worked until 1997 as Senior Consultant for the Norwegian computer security firm System Sikkerhet AS (now: Secode Norway), from 1997 to 2002 as scientist at the Norwegian Defence Research Establishment (FFI), from 2002 to 2004 as Senior Lecturer in information security at the Bodø Graduate School of Business, after which he took up the position of research scientist at SINTEF ICT. Mr. Jaatun is an expert in computer and communications security, electronic privacy, security evaluation, and system development. Mr. Jaatun served as security activity leader in WP2 of the IST OBAN project and was project manager of the Incident Response Management (IRMA) project funded by the Norwegian Research Council. Mr. Jaatun is vice chairman of the Cloud Computing Association.