Securoty and Privacy in a Structured Information Network

Tuomas Valtonen  
_Turku School of Economics and Business Administration, valtonen@utu.fi_

Tero Reuna  
_Turku School of Economics and Business Administration, tero.reuna@utu.fi_

Kalle Luhtitnen  
_Turku School of Economics and Business Administration, kalle.luhtinen@utu.fi_

Follow this and additional works at: http://aisel.aisnet.org/ecis2004

Recommended Citation

http://aisel.aisnet.org/ecis2004/164

This material is brought to you by the European Conference on Information Systems (ECIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in ECIS 2004 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.
SECURITY AND PRIVACY IN A STRUCTURED INFORMATION NETWORK

Valtonen, Tuomas, Turku Center for Computer Science, Lemminkäisenkatu 14–18 A, 20520 Turku, Finland, tuomas.valtonen@utu.fi

Reuna, Tero, Turku School of Economics and Business Administration, Rehtorinpellonkatu 3, 20500 Turku, Finland, tero.reuna@utu.fi

Luhtinen, Kalle, Turku School of Economics and Business Administration, Rehtorinpellonkatu 3, 20500 Turku, Finland, kalle.luhtinen@utu.fi

Abstract

The lack of trust amongst consumers and concerns about disclosing personal information are commonly seen as a major impediment to the growth of e-commerce. This is a consequence of several concerns, e.g. the lack of confidence in contemporary technologies for ensuring security, such as encryption and digital signatures, and distrust regarding the intentions of counterparts. This article describes the different aspects of security and privacy in the contemporary Internet and seeks means for improving both security and privacy. After identifying the fundamental components of Internet security and privacy, a novel software architecture aimed at enhancing these properties is presented. In order to evaluate the architecture, a two-dimensional model for expressing the extent of control in a network is introduced, and the security and privacy exhibited by the architecture is weighed against those in the conventional Internet. The consequences of increasing control in the Internet are also discussed. We conclude that although a new architecture encompassing greater control could be useful in certain areas of the Internet, a traditional uncontrolled area is essential to maintain unconstrained growth and the flexibility that has brought the Internet to where it is today.

Keywords: Internet, security, privacy, control.
1 INTRODUCTION

Advances in information and communication technology (ICT), during the past decade in particular, have changed the industrialized world significantly. The rapid expansion and commercialization of the Internet is one of the most notable phenomena of this period. The number of terminals connected to the Internet was estimated to have risen to almost 150 million in 2002 (Tietoyhteiskunta – Communications 2003). From a corporate point of view, the growth of the Internet community entails greater sales potential, as the Internet becomes an increasingly significant mediator between consumers and corporations. Although vast expectations were placed upon electronic commerce (e-commerce), a number of start-ups have been compelled to cease operations, and the market value of many of those remaining has dropped considerably. (Chang, Jackson et al. 2002, 1.)

The lack of trust amongst consumers and concerns about disclosing personal information are commonly seen as a major impediment to the growth of e-commerce. Although the number of online consumers is constantly increasing, a significant portion is reluctant to use services involving online payment. (Luo 2002, 1.) This lack of trust is a consequence of several concerns, e.g. the lack of confidence in contemporary technologies for ensuring security, such as encryption and digital signatures, and distrust regarding the intentions of counterparts (Belanger, Hiller et al. 2002, 249). Advances in technology are also opening new opportunities for harmful and malicious applications. In order to establish adequate protection against this threat, corporations and individuals are compelled to commit both time and financial resources. (Duke 2002, 4–5.)

Privacy and security play a central role in building trust in the Internet environment. Concerns about privacy amongst Internet users are often associated with the gathering of personal information, the means of exploitation and access to such information. (Udo 2001, 165.) Sheehan and Gleason (2001, 34) refer to a study, according to which consumers are less concerned about their privacy when the counterpart organisation is considered trustworthy. In the Internet environment this is problematic because the counterpart may be totally unknown to the consumer and situated in another continent. In such cases, risks related to disclosing personal information are often considered too high for the transaction to take place. Because many electronic services require sensitive information from the consumer, e.g. name, address, phone number and credit card number, many potential transactions are hindered by the reluctance of consumers to disclose mandatory personal information. (Hoffman, Novak et al. 1999, 82.) As a consequence, the lack of trust between the different parties operating in the Internet is restraining the overall growth of electronic commerce (Suh & Han 2002, 2).

In the Internet security is required, for example, in order to protect sensitive consumer information. In the absence of adequate security an intruder may gain unauthorized access to a system and inflict damage both upon the organization and the consumers that have disclosed personal information to the organization. Such actions could be motivated by the desire to damage the image of the targeted organization, to sell any acquired information to third parties or to exploit such information within operations of the intruding party itself. Although such threats cannot be completely eliminated, the risk of intrusion can albeit be reduced to a tolerable level. (Sanderson & Forcht 1996, 32.)

In this article, we first discuss the concepts of security and privacy in the Internet. We then present a novel software architecture that could complement the contemporary Internet technology. We also introduce a two-dimensional model for expressing the extent of control in a network, and weigh the security and privacy exhibited by the architecture against those in the conventional Internet. Finally, we discuss the possible consequences of deploying the new architecture at a global level.

2 SECURITY AND PRIVACY IN TODAY’S INTERNET

As mentioned in the previous section, the sense of insecurity amongst Internet users is a major obstacle for building trust towards the Internet, which is essential for the long-term growth of electronic
commerce. In this section we shall elaborate on security and privacy in today’s Internet and discuss the challenges faced when striving to guarantee security and privacy using contemporary technology.

2.1 Security

The Internet has created a new realm of opportunities for illegal activity. Because one can act anonymously or assume false identities in the Internet, the environment allows deceptive undertakings, such as fraudulent requisitions from electronic stores (e-stores), gathering of sensitive information via unauthorized electronic mail (e-mail) monitoring, the disruption of a web site by sending viruses or the illegitimate acquisition of credit cards numbers. Such illicit actions – and the threat of these – entail increased security costs and consequently higher product prices, and are thus harmful for both organizations operating in electronic commerce and consumers. (Laudon & Traver 2002, 231–232.)

The amount of financial loss accountable to illegal activities in the Internet is difficult to estimate because many organizations are reluctant to report incidents in the fear of losing credibility (Post & Kagan 2000, 14). Even in cases where an illicit incident is reported, assessing the extent of financial loss is all but a straightforward task, due to many potential indirect impacts of the incident. Moreover, intruders often pursue objectives of non-financial nature, such as attracting attention for ideological reasons, e.g. by disrupting the web service of an organization. (Laudon & Traver 2002, 232–233.)

The security of information is frequently seen as limited to that in computers and technology: protecting systems against hackers and viruses, or implementing backup policies for system information. However, security can be viewed as a broader concept with a variety of sub-disciplines. Below is one possible categorization for information security: (1) organizational security, (2) personnel security, (3) premises security, (4) information processing security, (5) telecommunication security, (6) ICT equipment security, (7) software security, (8) maintenance security, (9) storage media security, and (10) privacy of information (Järvinen 2002, 112–113). Due to the broad nature of the concept of information security, we focus in this article on categories 4–10.

2.2 Privacy

Advances in ICT, and the Internet in particular, have also created revolutionary methods for collecting and analyzing information. Each time that an Internet user visits a web site or follows a hyperlink from one web page to another, a trace of the user’s action can be stored in a log file (Strauss & Rogerson 2002, 174). Information is gathered both directly and indirectly. Direct gathering requires user action, such when as a consumer enters information in order to participate in an online competition. Indirect gathering may take place without the knowledge of the consumer, e.g. by using techniques such as cookies. (Steinke 2002, 194.) In either case, the type or extent of exploitation of the given information is often beyond the control of the consumer. Some organizations may attempt to collect all possible information about the consumer, even social security, credit card and bank account numbers (Sheehan & Gleeson 2001, 31).

Organizations are gathering enormous amounts of information, primarily intended for building new customer relationships and understanding the needs of consumers (Nakra 2001, 272). By using such information, individual consumers can be reached precisely and efficiently in comparison to traditional mass marketing. For instance, instead of displaying identical advertisements for each visitor at a web site, advertisements can be selected to best match the interests of each individual consumer. This process of consumer profiling can also be apply to enhancing the targeting of conventional printed advertisements, by using the real-world address of the consumer. (Graeff & Harmon 2002, 304.)

However, consumers are not indifferent to the means of collecting personal information. According to a survey by Hoffman, Novak et al. (1999, 81–82), the attitude of consumers towards traditional methods of information gathering is far more positive than towards Internet-related techniques. Up to 87 percent of the respondents expressed a need for more precise control over personal information gath-
ered via the Internet. Moreover, 21 percent were willing to receive advertisements via postal delivery, in contrast to only six percent via e-mail.

The Internet has also spawned a number of information brokerage organizations that gather information from cooperative organizations, accumulating huge amounts of personal information on Internet users. Many of these organizations also collect information from various public databases, such as telephone directory listings; some have also gained unauthorized access to confidential information, e.g. medical and financial records. Organizations in this category are prepared to sell any information in their possession to third parties, and thus represent a significant threat towards the privacy of Internet users. The falling of such information into the hands of a malicious individual or organization may result in grave consequences, such as blackmail. (Prabhaker 2000, 163.) This in turn could have a major negative impact on the behaviour of consumers in the Internet (Luo 2002, 112). The processing of personal information is regulated by legislation (e.g. Directive 95/46/EC of... 1995).

3 A STRUCTURED INFORMATION NETWORK ARCHITECTURE

In this section, we propose an overall architecture for organizing, distributing and personalizing content in the Internet. The Structured Information Network (SINE) architecture exploits the existing Internet infrastructure, i.e. the communication network and the terminals, supplemented by functionality for the administration, classification and filtering of content. The SINE architecture is constructed by deploying and configuring SINE core software; no programming is necessary in the network.

With respect to security and privacy, the main objectives for SINE are: (1) control over physical entities in the Internet, such as users, content providers, terminals and the network; (2) control over logical entities, i.e. the content and services in the Internet. In the following subsections we first introduce the technological structure of SINE and describe the different roles that physical entities can assume. We then elaborate on the mechanisms for implementing control over both physical and logical entities.

3.1 Technological Overview

The SINE architecture facilitates the flow of organized generic content from any content provider to any consumer terminal in a controlled manner via a number of network nodes. Generic content is organized and processed by service providers that may create content hierarchies, functionality or completely new content from existing content sources. Content can be cached and filtered at one or more intermediary network nodes along the path from the content provider to the consumer. The high-level data flow in SINE is illustrated in Figure 1.

![Figure 1. High-level data flow in SINE.](image)

Hence, we can identify three main roles for physical entities in the architecture: (1) content providers, (2) service providers and (3) consumers. The content provider creates generic content that is then made available to other parties in the Internet. The generic content could be any type of digital information (such as text, still image, audio and live image) or any combination of these, i.e. multimedia. Content providers must also publish a content hierarchy containing all generic content that is made available to other parties, in addition to a specification for each type of generic content in the listing, describing the structure of the content type. The service provider exploits generic content created by
content providers. The service provision role can be further sub-divided into three categories: (1) or-
organization and classification of content, (2) processing and combining existing content, resulting in
new generic content and (3) creating functionality. The objective of organization and classification is
to create a semantic content hierarchy, allowing parties to efficiently locate content and understand the
relationships and dependencies between different content objects. Processing and combining content
entails gathering content from one or more sources and performing logical or numerical calculations
on the content. The result is a generic content type that can be used locally (e.g. in a personalized user
interface) or made available on the network for exploitation by other parties. Finally, creating func-
tionality involves defining control dependencies between content objects and external controls.

Besides the aforementioned primary roles, certain supporting roles are required in SINE. These in-
clude (1) network operators that maintain communication links between network nodes and (2) termi-
nal administrators that provide the computer equipment, in addition to auxiliary operating systems and
databases required by the SINE software.

One should note that a physical entity could assume more than one of these roles. For example, an or-
ganization may produce proprietary content and make the content available to others (content provider
role), organize the content to form one or more content hierarchies and services for users (service pro-
vider role) and possibly exploit third-party content to create more elaborate services (consumer role).

3.2 Control Over Physical Entities

Physical entities in SINE must be registered and authenticated before access to the network is granted.
Registration applies equally to (1) physical entities operating in different roles (content providers, ser-
vice providers and consumers) and (2) the network infrastructure (terminals and network nodes). Content
and service providers are registered in a central database that contains information on the types of
content that a content provider is allowed to produce and valid categories in hierarchies created by a
service provider. For example, a car retailer would be authorized to define content categories that are
related to automobiles, but not content related to on-line music or consumer electronics. All consumers
must also be registered in regional databases, which in turn are managed by a central authority.

Terminals and network nodes in SINE are also registered in a central database, in which a master table
of the network identification codes are stored. Hence, unauthorized terminals or networks cannot be
added without registration, or these will remain invisible to others. Network nodes can also effectively
filter network traffic by offering only specific content categories to its sub-network or by filtering con-
tent according to information embedded in the generic content.

3.3 Control Over Logical Entities

Similarly to physical entities, all logical entities, i.e. the content and services, in SINE must also be
registered. When creating a new type of generic content, the content provider must publish the specifi-
cation that indicates the exact structure of the content type. This specification, which actually is a form
of generic content itself, is stored centrally and made available upon access to the content type. Fur-
thermore, when a service provider organizes the content into a hierarchy, a classification category
must be defined for each content object, defining its relationship to the hierarchy. Mandatory registra-
tion of these categories allows efficient limitation of the types of information that a specific service
provider may offer to consumers. Service providers must also ensure that the content of a registered
category actually matches its definition, or face the risk of being removed from the global service pro-
vider hierarchy by a central authority.

Individual consumers are granted access to content according to their personal access profiles; only
content and categories for which the consumer possesses sufficient access rights will be visible to the
consumer. Hence, when combined with an online payment service, SINE could be configured e.g. to
provide commercial access-restricted Internet services.
4 ENSURING SECURITY AND PRIVACY

In the previous section, we elaborated on SINE, in which the level of control is significantly higher than in the contemporary Internet. In this section, we discuss the impacts of imposing control over the physical and logical entities of a network. We first introduce a model for expressing the extent of control in a network, and position some exemplary networks within the model. We then view the architecture in the light of the challenges for security and privacy mentioned in Section 2.

4.1 The Network Control Model: the Internet vs. SINE

In Section 3 we distinguished between two aspects that express the extent of control in a network: (1) control over physical entities and (2) control over logical entities. Obviously, these two measures are independent of one another. For example, a network could have very stringent restrictions on network access, but unlimited access to data and applications. On the other hand, access to certain data and applications may be tightly controlled in a public network. Hence, we shall view these as separate measures and explore the combinations of these in a two-dimensional model, shown in Figure 2.

The contemporary Internet is an illustrative example of a network where most physical and logical elements are virtually unrestricted. Although Internet Protocol (IP) addressing schemes are controlled to a certain extent by Internet service providers (ISPs), the IP addresses that identify computers can be modified to a certain extent by users. The address of the sender of an e-mail can typically also be determined by the user without restraint. Today some portable computers even allow the media access (MAC) layer address, which should uniquely identify the network adapter, to be set by the user. Internet users are also allowed to operate completely anonymously in most areas of the Internet. Users can operate behind anonymous e-mail addresses and set up web sites\(^1\) without being identified.

There is also little control over the content in today’s Internet. A user with access to a web site can publish any material at his or her disposal. Although legislative authorities or ISPs can potentially monitor the content published at web sites, most Internet traffic remains uncontrolled. Furthermore, there is often no means for verifying the origin or correctness of content published in the Internet.

SINE, on the contrary, represents a strictly controlled network, both for physical and logical entities as described in Section 3. Physical entities operating in all roles (content providers, service providers, consumers) must be registered and authenticated. In addition, elements of the network infrastructure (terminals and network nodes) must also be registered in order to join the network. Logical elements (generic content and services) are also registered and controlled. The contemporary Internet, a partial SINE architecture with control over physical entities only (SINE-P), a partial SINE architecture with control over logical entities only (SINE-L) and the full SINE architecture are illustrated in Figure 2.

In Section 2.1 we listed 10 aspects of information security that should be taken into account in network design. Of these, we shall consider the latter seven categories: information processing security, telecommunication security, ICT equipment security, software security, maintenance security, storage media security, and privacy of information.

When only physical entities are controlled (SINE-P), information processing security, software security, storage media security and the privacy of information are uncontrolled. Hence, information security in these categories depends entirely on the selection, configuration and integrity of operating systems, applications and storage media, in addition to the information privacy policy of the organization. The security of telecommunication, ICT equipment and maintenance is ensured by allowing only authorized access (both normal users and maintenance personnel) to the necessary resources.

When only logical entities are controlled (SINE-L), the security of information processing and software depends on (a) the reliability of the service provider that created the functionality and (b) the storage media security. In criteria (a) we assume a service provider registered by the central authority to be reliable; if this is not the case, it should be removed from the network. Storage media security in SINE is equivalent to the database security, in which the information is stored. Assuming that database access is managed correctly and that the integrity of database is uncompromised, the storage media is secure. Privacy of information is ensured, because only those content categories that a user has access to are visible to the user; here we assume that administration manages the access rights correctly.

Hence, when both physical and logical entities are controlled security is ensured in all seven categories of information security (including “privacy of information”; i.e., privacy is implied by security).

5 DISCUSSION

The impacts of deploying SINE at a global scale could vary significantly, according to the commercial and political climate in different regions. At best, SINE could provide consumers secure access to reliable content on the Internet. Content providers could create, register and publish generic content on the Internet. Service providers could organize content from a number of sources, combine these and include functionality – creating complete services for consumers. Consumers could select from a variety of well-organized content and services. Due to effective control over both physical and logical entities, the network would be secure and the content reliable. As a result, consumer confidence towards the network could rise, spawning new growth in the Internet and e-commerce industries.

On the other hand, limiting access to the network could hinder the expansion of the Internet. Applying restrictions to content provision and service development could increase start-up bureaucracy, thus giving larger corporation a disproportionate advantage. Restrictions would also be in contradiction with the basic nature of the Internet: an open, unrestricted environment allowing flexible expansion and unconstrained creativity. Anonymity is also regarded a significant benefit for users of the Internet, especially when dealing with sensitive issues. Mandatory registration of all users could fuel suspicions of being monitored and personal information being collected. In some regions, the deployment of SINE could lead to the limitation of the freedom of speech; citizens would no longer be able to express their opinions under the cover of anonymity.

We conclude that SINE could be useful in areas of the Internet where high security and privacy is required, but the traditional Internet should be left intact areas that require flexibility and anonymity.

6 CONCLUSIONS

In this article, we first describe the significance of security and privacy for consumers operating in the Internet. In particular, we distinguish between the two terms, which are often considered a single con-
cept. We then elaborate on security in the Internet, classifying different aspects of information security. We discuss the threat towards the privacy of personal information in the Internet, as numerous organizations are gathering ever more detailed information on consumers.

We then propose SINE, a novel software architecture, to complement the contemporary Internet. We first describe the technical structure of the architecture and the roles of different parties operating in SINE. We then depict the mechanisms for controlling access to physical and logical resources within the network. Next we present a two-dimensional model for expressing the extent of control in the network, and position the Internet and variations of SINE in this model. We also discuss the impacts that could occur if SINE was deployed at a global level, and conclude that although SINE could be useful in some areas of the Internet, the traditional uncontrolled Internet is essential in others.

References


