Making Security Manifest

Security and Autonomy for End Users

Allan Friedman and L Jean Camp
{allan_friedman, jean_camp}@ksg.harvard.edu
79 JFK Street
Harvard
Cambridge, MA 02138

Manifest: Clearly apparent to the sight or understanding; obvious.

Abstract

With the increased concern over national security there has been increased debate over reliability and security of communications and computer systems. (Bush, 2001). One element of this effort has been on the need for reliable data on computer security risks and incidents. Information is necessary for a functioning market, and transparency (requiring information) is necessary for functional governance.

Despite the consensus on the need for better information, there is a significant divergence about the nature and distribution of security information. Security infrastructures can be mechanisms of user control (Anderson, 2003) or mechanisms to empower the user (Clark and Blumenthal, 2000).

End user security is critical. Distributed denial of service attacks illustrate how the capacity to create zombies (machines under the control of a malicious external agent) at many small nodes creates risks for the largest and most hardened targets.

Proposals to address failures in the market for computer security include the establishment of a liability regime for computer security, insurance markets for security risks or even creating tradable permits (Camp & Wolfram, 2000). Yet the solutions assume that the end user will be able to avail themselves of the legal or commercial mechanisms for security with little or no cost.

For security to function at the end points, there must be reliable data for the end user. Such data must be communicated clearly and there must be feasible mechanisms for the end user to respond to security breaches. In this work we illustrate that all the necessary technical components exist. What is needed is a vision and a national commitment to integrate the components. Developing the system that enables Internet users to protect themselves is a more powerful, more democratic and more resilient system for protecting our national information assets.

Summary

Information is necessary for a functioning market. Yet currently, security information is held by specialists, and no simple means exist to transfer that expertise to the average user. Secure systems are luxury goods, available only to those with the most discerning clientele capable of distinguishing, or for those businesses the high cost of failure justifies large security investments.

(Dean, 2003). In this paper we describe a mechanism to make it possible for a consumer to easily view the security of a remote system or their own.

Enabling ubiquitous, but not intrusive, security reporting can allow the consumer market for security to function. We refer to this as "manifest security." We propose three layers for the creation of a security market.

- First there is the specification of a XML scheme to allow for extensible expressions of trust as these develop on-line over time.
- Second, there is the reputation or recommending layer. This requires the capacity to use collaborative filtering or simple rating/reputation systems. This consists of the participants and selection of participants.
- Third, there is the reporting layer. In the first implementation, the reporting layer would be a browser applet, or a stand alone display (Yurcik et al, 2003).

Sufficient market pressures or other incentives are necessary to enable the diffusion of this overlay infrastructure. Yet there must be a proposal and proof of concept to enable adoption.

Motivation

That security is a market failure is the subject of some agreement.. A commonly proposed solution to rectify this market failure is the establishment of a liability regime for computer security (Anderson 94). Other potential remedies for market failures in security include creating tradable permits (Camp & Wolfram, 2000) and insurance markets for business risks (Schnieier 2002, Blakely 2002). Most of the discussion, however, centers on the enterprise and organizational actors, the traditional consumer or high-end security goods. Yet security is also a key factor in consumer behavior online, remaining a chief concern for online consumers (Deringer 2002) Shared information has the potential to transform security from a rare luxury good to a common commodity.

Reputation markets have a result in terms of price. While it is intuitively obvious that reliable merchants can charge more than unreliable merchants; it is also supported by a game theoretic approach to bidding in the face of reputation information (Dellarocas 2003), as well as empirical evidence from online auctions (Resnick & Zeckhauser 2002). We assert that security is another facet of reliability in the larger context of electronic interactions, exposing the actors to failure of not only the transaction, but potentially their entire system and loss of valuable information.

The majority of security literature focuses on the administrator level, preventing attacks on the technical side with a level of expertise outside the range of most non-technically trained users. If security is salient to their user experience, some accessible user-friendly mechanism should allow users to accept some responsibility for the security of their online interactions.

Trust Systems

In terms of effects on trust in computers and computer-mediated activity and readiness to accept security failures and move on, naïve users are not given the information to allow them to discriminate on the basis of the origins of harms such as memory damage, denial of service, leakage of confidential information, etc. In particular, it does not matter whether the harms are believed by users to be the result of technical failure, on the one hand, or human (or institutional) malevolence.

For example, key revocation policies and software patches all have an assumption of uniform technical failure. A key may be revoked because of a flawed initial presentation of the attribute, a change in the state of an attribute, or a technical failure. Currently key revocation lists are monolithic documents where the responsibility is upon the key recipient to check. Often, the key revocation lists only the date of revocation and the key. The social sciences would argue that the three cases listed above would be very different and would be treated differently. Consideration of that possibility leads to a key revocation system that may better fit human consideration of trust, and may manage risk more effectively, as well.

What about the case of an incorrect initial attribute? In this case, the possibility of malevolent action is most likely. Consider identity theft, since identity is a favored attribute linked to public keys (and was required by the first X.509 standard). Identity theft would call for more than revocation at the date of discovery. In a web of trust system; for example, should the revocation be narrowcast to anyone whose key or reputation is authenticated by the stolen identity? Any extension of cumulative trust enabled by the use of the key should be removed, and this should occur recursively until the entire result of the stolen identity is removed. Alternatively any account set up or configured with this key should be terminated. Ensuring that this domino effect of trust reversal occurs depends on user action (checking keys) as well as key lifetime. If users extend trust too aggressively without differentiating then a shorter key lifetime is required than if users increasingly differentiate risks. The capacity to create additional accounts and thus implement a domino of trust extensions is exactly the feature that makes identity theft attractive. Thus, key revocation schemes should take into account this capacity when evaluating methods for addressing the revocation of a particular key, and inform users who inquire about the key.

Consider a change in the state of an attribute. For example, a particular employee may be unauthorized to charge a company account after a sudden, unexpected, or particularly unpleasant termination. In this case, again, accounts that may have been created for the duration of the certification should be reconfigured. An example may be an account at a B2B exchange that requires certification at account initiation and considers the key lifetime, as set by the employer, as the appropriate duration of a valid account. Noting that this is a sub-optimal policy by the exchange is not likely to prevent flawed policies from being adopted; in particular when the interest of the businesses and the exchange is to accept risk in order to prevent denial of service. Recall that the Electronic Funds Transfer Act was initiated by exactly this type of change in attribute and malevolence, although in that case the malevolence resulted from divorce and not employment termination. The card issuer had a policy that expected individuals to know in advance how long the attributes -- in that case the marriage -- would last. Learning from past failures in payment systems depends on learning about the trust failures in these payment systems. In contrast, given a technical failure of a lost key, all that would be necessary is preventing future assertions by the holder of the subverted key. By having a single standard key, revocation systems implement the assumption that there is no significant systematic difference in people's reactions to betrayals that originate from human actions, on the one hand, and computer failure, on the other.

With respect to software patches, the possibility of a purposefully malevolent alteration of the code is not considered. The social sciences would argue that such cases require a different level

of active response and oversight than technical error made in the market-equivalent of good faith. For example, a bug purposefully placed by hackers who had access to Microsoft's source code would presumably be meant for harm; while the other 63,000 bugs in Win2k could be considered either minor or less likely to enable malevolent action. Thus a discovery of a malevolent bug should result in active contact with all customers who had installed the product and technical support to enable effective patching; while the standard policy of customer-driven downloading of patches could be adequate for other cases. Yet the users must not be forced to upgrade at the moment the vendor chooses. Currently customers do not know of the existence of bugs, and the proposed practice is forced upgrade. There is an obvious middle choice being ignored in the technical and policy debates.

Consider how people distinguish between computers. In terms of best practices for security, it makes most sense for people to view distinct remote computers as distinct individuals, each one warranting independent evaluation. Yet, there are several lines of research that converge at quite a different point, suggesting that users tend to view networked computers as constituting a more homogeneous system. Users have no been provided adequate information to manage the security of their own systems. An analogy would the car with a drive train but no steering wheel or displays of automobile status. Centralized remote driving would not be the obvious solution to such a situation.

Descriptive Mechanism

Obviously, the development of a flexible, extensible XML lexicon describing a full range of security issues is a non-trivial task. Development should involve a full range of interested parties, and adhere to a fully open, participatory process. Where possible, the scheme should reflect easily verifiable information that can cleanly be stated to be either true or false (i.e. is the latest patch installed). Grammatical structures should reflect the ever-evolving nature of security requirements and be as open as possible. The development process of W3C's Platform for Privacy Preferences can be held as a model for cooperative and effective descriptive language design. There are multiple efforts to create full XML descriptions of security, in particular the Security Assertion Mark-up Language. In addition Fiegenbaum and Blaze have developed a language for expressing trust policies.

Reputation and Recommender Systems

If individual users lack the expertise and resources to assess the security of the systems with which they interact, then the next best solution is to rely on trusted experts. Of course, it is never so simple: one seldom knows whether to trust an expert. Information economics has recently seen a rich development in the field of recommender or reputation systems as a means of determining trustworthiness. Recommender systems make up the backbone of product recommendation mechanisms (i.e. Amazon); they also serve as crucial social norms enforcers in public forums (i.e. Slashdot). At its heart, such a system consists of some means of allowing some subset of a community to voice their opinions on the value of the agent or product being evaluated; this feedback is aggregated and used by other actors in their relationship to the object of the feedback. Such mechanisms can create records for developing trust in a society of strangers (Dingledine 1999; Camp 2003).

We can imagine expertise varying across a range of dimensions, each defined in terms of a security concern of a user. For instance, a reader of an online newspaper may want to know that the site has not been defaced; a user of chat rooms may want to know that her anonymity is being protected; a consumer of health information would care about the security of personal information collected. Someone shopping on the ACLU website might care about all three. Along each of these dimensions, a security analysis of a given system can be offered by any party. Other actors can support or take exception to this evaluation, and the reputation system can aggregate these sentiments.

Each user may choose to select a reputation provider. The obvious candidates for reputation systems include incident response teams, software vendors, and security consultants. Incident response teams are interested in increasing network security. Software vendors can offer high ratings to companies that download and install patches in a timely manner. Security consultants can offer their expertise to a broader market.

While in the privacy field, third party verification in the form of Privacy Seals has met with limited success, we believe that by balancing third party reputations with the feedback of a recommender system, a useful measure of aggregated trust can be established in a security evaluation system. The problem of incentives of recommenders becomes a problem.. Unlike a book or an auction seller, refutation of a positive recommendation (or, to a lesser extent, a negative comment) will not be immediately confirmed for feedback into the system. Similarly, unlike a purchase there may be no direct cost for a particular entity to create a recommendation.

While widespread participation creates its own set of risks, having a large number of users in addition to a small set of experts is optimal. Increasing the number of observers increases the probability that an action is observed. For example, many viewers increase the chance of viewing a web page defacement, receiving notification of a need for a new credit card, or experiencing fraud directly connected to a subverted server. However, the risk of a security failure of any one system is relatively low; we thus cannot depend solely on a self-correcting reputation system for security analysis. Trusted third parties can claim the authority to evaluate the expertise of evaluators, inspiring trust, or assume the role themselves.

The system may collect ratings with simple UDP requests that are sent in parallel with http requests. Alternatively the user could have a learning agent that trusts its own user assertion of positive and negative on-line experiences and collects information of others. In this case reputations for frequently visited pages would be updated by the agent, and the agent could build from the person's browsing history to check on the sites regularly visited.

Each system has its own benefits and drawbacks in inspiring trust. Optimally, some combination of external expert recommendations, reputation systems and logged personal preferences would enable a user to feel confident about the security reputation of the system with which she is interacting. A non-security analog would be a reader who uses the New York Times Book Review, epinions.com book rating system and amazon.com's personal recommendations page to select her next volume.

Displaying the Result

Once a level of security has been determined, the user needs a relatively simple and real time interface. We have no reason to believe that users will undertake this research themselves, and thus claim that an unobtrusive interactive agent is needed to deliver security information for effective use. A user will need to interact with the security system in two ways.

First, she will need to configure the system to reflect her preferences in terms of security and the parties she trusts to evaluate the security of the internet hosts with which she will be interacting.

Second, she will need to learn whether or not a specific action conforms to these preferences. Three possible models present themselves for feedback mechanisms.

First, the user can be alerted by a simple Boolean trust-worthy / not-trustworthy notification. The lock icon for the Secure Sockets Layer offers this model for communications confidentiality. Since security is hardly such a one-dimensional variable, the user should have some idea of what being trust-worthy entails, either by understanding what her trusted agent is examining, or having expressed her security preferences to an agent that interprets the security evaluation from the reputation mechanism. There is increased work on reputation systems, which allow for trustworthy users and sites to be identifier through cooperative information sharing. (Dingledine 1999; Camp 2003).

Second, the multi-dimensional aspects of security can be converted into a one-dimensional scale, expressed as a number in a given range or a color in a given spectrum. This also implies some awareness of what more secure or less secure means to the user, but allows gravitation towards a greater degree of security. Finally, a multi-dimensional feedback mechanism, perhaps in the presence or absence of small icons, can alert the user to various levels of security in terms of, for example, encryption and back-end database protection. After a while, a user will learn to submit her own security recommendations or experiences as a rating into the reputation mechanism.

Two models we use as inspiration are AT&T's Privacy Bird using W3C's P3P standard, and William Yurcik's NVisionIP network security tools. The Privacy Bird, a browser plug-in that employs visual and auditory cues to keep the user informed about the privacy promised by a viewed web page. The Privacy Bird allows users to express their privacy preferences in easy to understand language, and also provides even simpler abstractions of 'low,' 'medium,' and 'high' levels of privacy. It can be positioned around the screen and its visual and noise cues can be set for maximum convenience and minimum annoyance.

The NVisionIP security tools are specially design to illustrate security state although not for the most naïve user. (Yurcik et. al., 2003) The NVisionIP security tools, developed at NCSA, allow users to easily interpret a vast amount of network traffic administration with visual pattern and color information. Additional interpretation or complexity setting may be required, yet there is no reason a home user should have such difficulty detecting that their own machine is sending out SPAM or scanning others.

The two major critiques of the Privacy Bird in a recent study are a lack of P3P implementation across the web and an inability to trust the privacy policies available (Cranor et al 2002). The latter is addressed in a manifest security model by a transfer of trust to a consumer-selected

trusted entity; the issue of implementation will be addressed below. The key features desired in this system of accessibility and multiple levels of customizability offer easy-to-use security information and allow for more responsible and informed internet use.

Potential Market Configurations

This paper offers three models under which manifest security might be propagated across a large segment of the internet-using non-expert population: fear of attack, de facto or de jure regulation and the more market-centered liability expansion.

As computer security becomes a more mainstream topic in the media, and computer security threats rise, it is possible that a tipping point could be reached in terms of mass perception of security exigencies. Several high-profile security debacles, for example, might make it easier to conceive of a personal security threat. In this case, public demand should drive a market solution to some security issue. Users would actively seek reporting clients, and the services of commercial or non-profit security evaluation agencies, if they were available. Manifest security offers a possible solution that is less extreme than a wholesale shift to secure systems such as TCPA or Palladium, and allow greater user autonomy.

Alternatively, browsers could start being shipped enabled with some manifest security technology built in and turned on. This could come from a government mandate for some national security policy or, as was the case with P3P, acceptance of the user-end technology by the market leader in browsers. Without a sufficient demand, however, it will be much more difficult to build the reputation infrastructure, and the technology might languish as a not-wholly assimilated system. The largest browser provider has an incentive to provide ratings that underestimate the security of competitor's products.

Finally, we can speculate about an environment in which the liability standards that have been advocated for quite some time are indeed adopted. If parties responsible for security-related damages are held accountable, it is likely that some of this accountability will filter down to the end user. Machines hooked up to always-on, high-speed internet access, for example, have recently been implicated in the rise of Distributed Denial of Service attacks. If the victims of these attacks hold the ISP's liable for damages, ISPs would thus have a very strong incentive to encourage proper security practices in their customers. They would encourage both the use of interactive agents and the acceptance and reputability of trusted security experts, strengthening both sides of the system.

Each of these scenarios can be analyzed to speculate what a large-scale implementation of manifest security might look like. This research is currently in progress.

Discussion and Benefits

Under this guise, security is taken as a super-set of currently applied observable values: is the latest patch installed, did the latest audit find leaking data or is this site using an easily breakable key. (For example, a survey of keys in 2001 found keys as short as 40 bits, see van Somerson 2001). Manifest security is the means by which a user of information systems can be sure that information being sent and received is valid and can be trusted. Market forces develop when demand awareness are high enough to place consumer value on security information; that

information could be paid for by the consumer seeking security, or a remote host seeking to accrue a reputation in order to earn the trust of customers.

There is a recognized need to inform users about privacy to enable them to make rational information management decisions. Similarly. more responsible, informed views for better decision-making is needed for information security. Making security manifest to the user will generate more public awareness of security, increase information, and thereby enable the security market to function in a way as to serve consumer interest.

Conclusion

The debate about securing the network has focused heavily upon created "trusted systems' where the owner of a computer becomes an operator while the owner of the content purchased by the computer user become the "owner" of the computer. (Anderson, 2003) The risks to users and the network in a centralized proprietary system are described in full elsewhere.

There has been an embedded assumption in the debate that users need to be controlled in order to secure systems.

In this paper we bring together the economics of security proposals, reputation systems, the work in HCI, and examples of using information to create effective governance to propose a system for ubiquitous security information for the end user. Such a system could enable end users to protect the network by preventing their own resources from being utilized in attacks. The long term risks to security, as well as freedom and innovation, are less in an end-to-end system than in a homogenous system with centralized control.

Anderson (2003), "Cryptography and Competition Policy – Issues with Trusted Computing", *Second Workshop on economics of Security*, May 2003. (College Park, MD).

Bush (2001), "Executive Order 12321 on Critical Infrastructure Protection"http://www.whitehouse.gov/news/releases/2001/01/20011016-12.html

Lorrie Faith Cranor, Manjula Arjula, and Praveen Guduru. Use of a P3P User Agent by Early Adopters. In Proceedings of the ACM Workshop on Privacy in the Electronic Society, November 21, 2002.

L. Jean Camp, ""Peer to Peer Systems", The Internet Encyclopedia ed. Hossein Bidgoli, John Wiley & Sons, (Hoboken, New Jersey) 2003.

L. Jean Camp & Catherine Wolfram, "Pricing Security" , Proceedings of the CERT Information Survivability Workshop, Boston, MA Oct. 24-26, 2000, pp. 31-39.

D. Clark & M. Blumenthal, (2000) "Rethinking the design of the Internet: The end to end arguments vs. the brave new world", *Telecommunications Policy Research Conference*, Washington DC, September.

D. Dean, "The Economics of Security", Financial Cryptography, Jan 29 – Feb 2, 2003. Guadeloupe, France.

Chris Dellorcas, Efficiency and Robustness of eBay-like Online Reputation Mechanisms in Environments with Moral Hazard, Center for eBusiness Working Paper #170, 2003

Deringer Research Group, American Interactive Consumer Survey, 2002

R. Dingledine, “Chapter 16: Accountability” . *Open codes: Voices from the Open code Revolution*, DeBona, S. Ockman and M. Stone (eds) O'Reilly, 1999.

Paul Resnick, and Richard Zeckhauser. “Trust Among Strangers in Internet Transactions: Empirical Analysis of eBay's Reputation System.” The Economics of the Internet and E-Commerce. Michael R. Baye, ed. Amsterdam, Elsevier Science, 2002.

N. van Somerson, “Key Lengths and Jurisdictions”, Financial Cryptography, Bermuda BWI, January 2001.

William Yurcik, James Barlow, Kiran Lakkaraju and Mike Haberman (2003), “Two Visual Computer Network Security Monitoring Tools Incorporating Operator Interface Requirements” Workshop on Human-Computer Interaction and Security Systems, (Florida).