Quantifying economic impact.

Not only are losses difficult to quantify with today’s tools, but it appears to be difficult to assess in economic terms the return on investments achieved via security measures. Measures of business and consumer investment in security-related activities lack a common reporting entity or information aggregating mechanism.

This is due in part because the return on security investment may accrue to an entity other than the investor. Just as “poor cyber ‘hygiene’ on one Internet-connected computer negatively impacts other connected computers, so does good hygiene reduces the risk of infection for all.

Yes, businesses obviously lack incentive to provide information about their failures and losses. Even without the (perhaps exaggerated) fear that its reputation will be damaged, the business has no incentive to create such reports, if it involves work beyond their normal processes.

Often, such studies are merely thinly veiled pitches for the security product of the author or his surrogate. It is not surprising that the conclusion of a computer crime survey funded by a vendor of security event management tools would find that such tools reduce the occurrence and impact of events. Even granting the truth of such self serving conclusion, such reports tend to be of limited value..

Perhaps one way to encourage a uniformity of reporting would be through the tax code. If cyber security incident related losses and investments in security related activities and infrastructure were specifically enumerated, it could establish a baseline of data that would begin to show useful trends. However, to the degree that tax policy provides incentives to report either cyber related losses or investments, such a policy would likely come with its own set of unintended consequences and skewing factors.

Political pollsters often site results from generic questions that have been asked year after year in exactly the same way, (e.g., “What party do you expect to support in the next congressional
election?’) because they provide useful trend information over time and serve as a sort of smoothing function that irons out anomalies. Cyber security data gathering should adopt a similar long term view. The kind of questions worth answering are “how quickly are things getting worse (or better)?” and “where are the primary areas of concern?”

2) Raising Awareness

Although the necessity of security awareness and training has been recognized for some time, its effectiveness as mitigation for risks outside of circumscribed environments has been overestimated. According to a July 2009 Network world article, the Zeus botnet includes 2.9 million computers. Even if a pervasive awareness program achieved the herculean task of convincing half those users to maintain their machines securely and avoid risky behavior sufficiently to escape being botted, Zeus’s botmasters would still be controlling a potent cyber weapon.

Heighted awareness provides little protection against “respectable” sources of infection. For example, in 2007, the Miami Dolphin's web site was compromised just before the Super Bowl and served malware to thousands of users. Since then, other sites managed by professional and presumably “aware” administrators have been similarly compromised. Nor does awareness reduce the number of hours people have to spend securing their machines. Windows, Adobe, RealPlayer, iTunes and dozens of other applications constantly nag users to install updates, and then require a re-boot. Even the security savvy can be forgiven if they choose forgo this cycle in favor of productive work occasionally.

One document sited by the ROI (National Initiative for Cybersecurity Education (NICE) Relationship to President’s Education Agenda 19 April 2010) states,

NICE will be conducting an aggressive nationwide awareness and outreach program (Track 1) designed to effect a cultural change in our society that will make the use of sound cyber practices when interacting with cyberspace as common as wearing seat belts when driving or riding in a car.

This is a worthy goal, and one can hope that such an initiative will have as dramatic an impact on cyber security as seat belts have had on auto safety. However, this analogy is only marginally applicable. First, seat belts come preinstalled in cars and their specifications are defined by federal regulation. Second, seat belts are simple to use, almost never wear out, and are hard to ignore, given the audible warnings that attend their remaining unbuckled. Third, the use of seat
belts is mandated by law nearly universally. Fourth, drivers who, in spite of all this, forgo their use risk injury to themselves, not to others. In contrast, robust computer security protections must protect against variable threats, not a single immutable law of physics, and their implementation requires constant vigilance and expertise. Additionally, unsafe computing practices represent a threat to the community, not just the heedless.

However, the automobile analogy may provide guidance. Automobiles are large, complex machines that must be properly maintained to provide their desired function and avoid damaging others. However, the level of knowledge and awareness demanded of the automobile user is rudimentary compared to the complexity of the machine. Drivers know that they must obtain oil changes and other scheduled maintenance, and they know that their cars must be periodically inspected. Cars whose brakes and lights don’t work are not allowed to endanger others. Conversely, computer users that merely want to read blogs, shop online, and use e-mail are routinely presented opaque warnings like, “The instruction at 0x7c901010 referenced memory at 0x0000001c. The memory could not be read.” Or they are confronted with messages importuning them to update this or that application, which typically entails the interruption of closing applications and rebooting. It is hardly surprising that people who prefer using their computers to maintaining them postpone such updates indefinitely.

In order to make an awareness program useful to the typical end user, the degree of complexity should be reduced to a level that is analogous to that which is required of the car owner. Some small steps are already taking place in this direction. The Google Chrome browser updates itself silently without user interaction, and the latest version of the Firefox browser offers this as a default setting that can be overridden.

Automatic updates should be default setting for all software. Knowledgeable administrators who need to test updates before they are deployed, and whose livelihood depend on maintaining secure systems can opt for a manual process. Typical home users using a vanilla configuration need not know that their applications are being silently updated.

Another way in which the inexpert user can be assisted is seen in the draft specification for Strict Transport Security, (STS) which defines a mechanism enabling “web sites to declare themselves accessible only via secure connections, and/or for users to be able to direct their user agent(s) to interact with given sites only over secure connections.” \(^1\) In short, STS will not allow the typical user to override baffling messages warning against expired certificates or “mixed content”. Inherent in this idea is a refreshing concept: websites that adopt it are effectively declaring a security covenant with their customers: their certificates will always be valid and properly configured, and if they are not, the website owner doesn’t deserve the public’s trust.

Underlying these recommendations is a security axiom that has been around for years: “default deny”. In the context of the issues discussed here, this adage might be better worded, “default secure”. That is, in any calculation concerning security vs. convenience, security should always win, in the absence of any specific requirement for a feature or permission.

---

3) Web Site and component Security

Web site security could be significantly enhanced by establishing a penetration testing process that is easy to implement, inexpensive, and which does not put private information at risk. Such a scheme might be implemented as follows:

a) The appropriate government agency would vet the members of a “white hat” hacker community, either by a third party certification process, a low cost background check, or both.

b) Business or other sponsors of public web sites would make a “pre-production” version of the web site accessible. The hardware and software of this version would mimic the production environment, but contain no actual private data.

c) Members of the vetted hacker community would be challenged to penetrate the security of the web site, and receive payment for succeeding.

This approach would have the following benefits.

a) It would allow companies to avail themselves of sophisticated security skills with very little risk, since the web site owners would pay “by the bug”.

b) Vulnerabilities discovered in commercial products could be made available in a controlled fashion, since the testers would be bound by a non-disclosure agreement, but product vendors would have an incentive to build secure software, since a scorecard of results could be released. Successful testers would be rewarded by monetary payment and the enhancement of their reputations as legitimate security researchers.

c) This approach would be adaptable to rapid technology changes and be compatible with industry product development and maintenance schedules and practices.

While there might be many details to work out in such an approach, and disputes among participants will inevitably arise, participants would have to agree to accept the arbitration of the sponsoring agency. Although the ethical acceptability of some “bug bounty” programs have been questioned, these concerns have mostly been leveled at commercial, third-party entities profiting from product flaws whose vendors are not willing participants in an established program. The focus of this effort would be identifying configuration and implementation flaws in established web sites with the active participation of their owners.

4) Authentication/ID Management

Is there a continuing need for limited revelation identity systems, or even anonymous identity processes and credentials? If so, what would be the potential benefits of wide-scale adoption of limited revelation identity systems or anonymous credentialing from a cybersecurity perspective? What would be the drawbacks?
We routinely offer a widely accepted credential (usually a drivers license) to facilitate in-person transactions, but there is no similar credential in the virtual world. Several years ago it was widely thought that Public Key systems would provide a basis for universal and robust identification for a host of on-line interactions.

However, the promise of PKI has not been realized. This is part a function of the herky-jerky nature of technical innovation, and partly because public key crypto was burdened with unreasonable expectations. The phrase, “non-repudiation” quickly became associated with PKI, suggesting that it could hold every certificate owner legally responsible for every keystroke, and that “wet” signatures heretofore necessary and sufficient for the signing of contracts would become a thing of the past. Much was written by lawyers about the the legal verifiability of digital signatures and many changes were made to the PKI standard by technologists to accommodate evolving needs.

Suffice it to say, the vision of PKI as the silver bullet of on-line identification never took off.

But if we were to scale back our expectations and requirements, PKI could serve some current needs admirably. Suppose you could go to your local post office, show a drivers license, pay a modest fee, and be issued a certificate that asserted only that on a certain date, the holder of this certificate identified himself to a post office in the state of Virginia. In other words, the certificate would assert only that the bearer is a real person with a real identity, not a spam-producer or a botnet. Businesses, large and small, represented by a real person, could do the same. Then, users could instruct their computers (or their ISP) to block all mail that wasn't accompanied by a “this is a real person” certificate. If a real person, who happens to be a spammer, obtains such a certificate, then it would be a simple matter to block the mail uniquely identified by that certificate. While the certificate would be cheap, it wouldn't be so cheap as to make “certified spam” a paying proposition. And if a certificate were associated with threats, harassment, or other malfeasance, a judicial warrant could be invoked to reveal the identity of its owner, using the same mechanisms that are available to law enforcement today.

This would also permit widespread use of encrypted email. Most email comes from repeat correspondents. With certificates all around, the first exchange of email would create a symmetric session key that could be relied on for a few weeks or longer.

A couple of possible objections:

W*wouldn't this destroy anonymity on the net, thereby undermining its usefulness as a medium of free expression and a useful tool for whistle blowers?

No. First off, a warrant would be required to identify the certificate holder. But even if one assumes that the judicial system is compromised, it is perfectly reasonable to suppose that anonymous proxies will accept mail, or other connection types without demanding a certificate. If the Washington Post, various hotlines, and private detective agencies and sites like Wikileaks choose to accepted non-certified mail, they will not make a tempting target for spammers or hackers, since they provide no conduit to the credulous souls that actually respond to spam or connectivity to systems worth subverting.
During the late 1980's there was considerable discussion about the privacy impacts of caller ID for the phone system. (http://catless.ncl.ac.uk/Risks/8.42.html#subj1.1). On balance, it would seem that most people prefer the advantage of knowing who is calling before they answer the phone.

** Wouldn't the expense and and hassle of maintaining the certificates outweigh the benefits? **

Given the low level of trust associated with the certificate, it seems unlikely. Potentially, the certificates might never expire. Certificates would only need to be verified occasionally, for example, when one received an email from a new correspondent. Mail to and from your bank, your doctor, your friends and the garage would all be encrypted with the symmetric key that was established at the first email exchange. Post offices are ubiquitous and have been in the business of facilitating the reliable delivery of messages for generations. It would be easy for certificate owners to solve problems and update keys when there is a walk-in help desk in every zip code.

** What if a user’s private key is compromised? Can’t he be impersonated? **

Illegally acquired private keys can’t be used for impersonation, since they don’t uniquely identify anyone. If a certificate is used to vet the online activities of a wrongdoer, it won't be long before the certificate winds up on users’ and ISPs' blacklists, Note, this isn't a CRL, at least not in the traditional sense, because it is distributed, and no entity is responsible for maintaining it. But when a user discovers that his mail is being rejected and his attempts to connect to sites that demand a certain level of trust are rejected, he will obtain a new certificate for a modest amount and probably ask himself, “How can I keep this one safe?”

Much research and effort has gone into the design of captchas to discriminate between automated and human agents. A modestly targeted certificate based system could accomplish the same goal and considerably more.