Economics, Psychology, and Sociology of Security

Andrew Odlyzko

Digital Technology Center, University of Minnesota,
499 Walter Library, 117 Pleasant St. SE,
Minneapolis, MN 55455, USA
odlyzko@umn.edu
http://www.dtc.umn.edu/∼odlyzko

Abstract. Security is not an isolated good, but just one component of a complicated economy. That imposes limitations on how effective it can be. The interactions of human society and human nature suggest that security will continue being applied as an afterthought. We will have to put up with the equivalent of bailing wire and chewing gum, and to live on the edge of intolerable frustration. However, that is not likely to to be a fatal impediment to the development and deployment of information technology. It will be most productive to think of security not as a way to provide ironclad protection, but the equivalent of speed bumps, decreasing the velocity and impact of electronic attacks to a level where other protection mechanisms can operate.

1 Introduction

This is an extended version of my remarks at the panel on "Economics of Security" at the Financial Cryptography 2003 Conference. It briefly outlines some of the seldom discussed-reasons security is and will continue to be hard to achieve.

Computer and communication security attract extensive press coverage, and are prominent in the minds of government and corporate decision makers. This is largely a result of the growing stream of actual attacks and discovered vulnerabilities, and of the increasing reliance of our society on its information and communication infrastructure. There are predictions and promises that soon the situation will change, and industry is going to deliver secure systems. Yet there is little visible change. Moreover, the same predictions and promises have been around for the last two decades, and they have not been fulfilled. At the same time, the world has not come to a grinding halt. Not only that, but, contrary to other predictions, there have not even been giant disasters, such as a failure of a bank, caused by information systems insecurity. The really massive financial disasters of the last few years, such as those at Enron, Long Term Capital Management, or WorldCom, owed nothing to inadequacy of information security systems. How can we explain this?

Growing ranks of observers have been arguing that one needs to understand the non-technical aspects of security, especially economic ones [1], [2], [4], [8],

[10]. Security does not come for free, and so it is necessary to look at the tradeoffs between costs and benefits. Furthermore, it is necessary to look at the incentives of various players, as many have an interest in passing on the costs of security to others, or of using security for purposes such as protecting monopolies. There is now even a series of workshops on Economics and Information Security (see [11] for information about the first one, including abstracts and complete papers). This note does not attempt to summarize the literature in this area. Instead, it briefly outlines some factors drawn from psychology and sociology that make security costly to implement, and thus require the economic tradeoffs that we observe being made. It also helps explain how we manage to live with insecure systems.

The basic problem of information security is that people and formal methods do not mix well. One can make the stronger claim that people and modern technology do not mix well in general. However, in many situations people do not have to be intimately involved with technology. If a combinatorial optimization expert finds a way to schedule airplanes to waste less time between flights, society will benefit from the greater efficiency that results. However, all that the passengers are likely to notice is that their fares are a bit lower, or that they can find more convenient connections. They do not have to know anything about the complicated algorithms that were used. Similarly, to drive over a bridge, all we need is an assurance that it is safe, and we do not require personal knowledge of the materials in the bridge. The fact that it took half as much steel to construct the bridge as it might have taken a century ago is irrelevant. We simply benefit from technology advances without having to be know much about them.

With security, unfortunately, technology can be isolated from people only up to a certain point. (The success of SSL/TLS was due to a large extent to its workings being hidden from users, so they did not have to do much to take advantage of it. That was an unusual situation, though.) As information technology permeates society, more and more people are involved. A system is only as secure as its weakest link, and in most cases people are the weak link. A widely circulated piece of Internet humor is the "Honor System Virus:"

> This virus works on the honor system.
>
> Please forward this message to everyone you know, then delete all the files on your hard disk.
>
> Thank you for your cooperation.

We can laugh at this, but in practice there are email messages popping up all the time, telling people that their computer has been infected by a virus, and telling them to find a file named "aol.exe" or something a bit more obscure, and to delete it. Moreover, a certain number of people do follow such instructions, and then plaintively call for help when they cannot connect to AOL. This is part of a pervasive phenomenon, in which people continue to lose money to the Nigerian 419 scam ("please help me transfer $36 million out of Liberia, and I will give you 20 percent"). Social engineering ("this is Joe from computer support,

we need your password to fix a bug in your computer") continues to be one of the most fruitful attack methods.

The standard response of technologists is to call for more and better education. However, that has not worked in the past, and is not likely to work in the future. Although education is useful, there will be countervailing tendencies (similar to those cited in [7]), namely more people will be using information and communication systems in the future, and those systems will be growing in complexity.

The message of this note is not that we should adopt a defeatist attitude to information security. The point is that we should be realistic about what can be accomplished. A productive comparison might be with auto safety. There has been substantial improvement in the past, and it is continuing. Greater crashworthiness of cars as well as better engineering of roads and more effective enforcement of drunk driving laws and more use of seat belts have made car travel far safer. In the United States, deaths per mile traveled by car fell at a compound annual rate of 4.7 percent between 1980 and 2000, by a cumulative factor of more than 2. However, because of growth in volume of travel (by 80 percent), the total number of deaths has only decreased from 50,000 per year to 42,000 per year. Moreover, in the last few years, the annual number of fatalities appears to have stabilized. Our society has decided (implicitly, without anyone ever voting on this explicitly) that we are willing to tolerate those 42 thousand deaths per year. Measures such as a drastic reduction in the speed limit, or devices that would constantly test the driver for sobriety or alertness, are not acceptable. Thus we manage to live with the limitation of the large masses of human drivers.

In information and communication technologies, we have also managed to live with insecurity. Chances are that we will manage to live quite well even with the projected insecurity of future systems. After all, we have lived without perfect security in the physical world. The locks on our house doors are not secure, nor are our electricity and water supply systems. In practice, though, existing safeguards are sufficient. Now the problem in cyberspace is that attacks can be mounted much faster and on a more massive scale than in the physical realm. The answer to that, though, is not to strive to build perfectly secure systems, as that is impossible. Instead, it should suffice to put in enough "speed bumps" to slow down attacks and keep their impact manageable. The reason this approach should work is the same one it has worked in the physical world, namely that it is not just the content of communication that matters, but also its context, and the economic, social, and psychological factors that hinder the deployment of secure systems provide protective mechanisms.

2 The Incompatibility of Formal Methods and Human Nature

A crippling problem for secure systems is that they would make it impossible for secretaries to forge their bosses' signatures. As was mentioned in [8], good secre-

taries know when it is safe to sign for their bosses to keep those bosses' workload manageable and speed the flow of work. There is some anectodal evidence that in organizations that move towards paperless offices, managers usually share their passwords with their secretaries, which destroys the presumed security of those systems. In a formal system, one can try to provide similar flexibility by building in delegation features. However, based on prior experience, it seems unlikely that one could achieve both security and acceptable flexibility.

In general, people like to have some slack in their lives. Sometimes this is exploited on purpose. Stuart Haber (private communication) reports that in marketing the digital time-stamping technology that he and Scott Stornetta invented, some accountants did raise concerns about losing the ability to backdate documents. As another example, when the U.S. Securities and Exchange Commission responded in 2002-2003 to the pressure to clean up corporate financial abuses, it attempted to make lawyers responsible for reporting malfeasance they encountered. The proposed wording of the rule had the definition [5]

> Evidence of a material violation means information that would lead an attorney reasonably to believe that a material violation has occurred, is occurring, or is about to occur.

However, lawyers objected to something this straightforward, and managed to replace it by

> Evidence of a material violation means credible evidence, based upon which it would be unreasonable, under the circumstances, for a prudent and competent attorney not to conclude that it is reasonably likely that a material violation has occurred, is ongoing, or is about to occur.

We can of course laugh at lawyers, but our lives are full of instances where we stretch the rules. Who does not feel aggrieved when they receive a speeding ticket for going 40 when the speed limit is 35?

There are also deeper sources of ambiguity in human lives. For example, suppose that you need to go to work, and you leave the key to your house or apartment with your neighbor, with the message "Please let in the plumber to fix the water heater." Seems like a very simple and well-defined request. However, suppose that after letting in the plumber, the neighbor sees the plumber letting in an electrician. You would surely expect your neighbor to accept this as a natural extension of your request. Suppose, though, that your neighbor then saw the plumber and the electrician carrying your furniture out. Surely you would expect the neighbor to call the police in such cases. Yet the request was simply "Please let in the plumber to fix the water heater." It did not say anything about calling the police. Human discourse is based on a shared culture, and the message "Please let in the plumber to fix the water heater" embodies expectations based on that culture. That is why we do not leave our keys with our neighbors' 6 year old daughter with such a message.

A particularly illustrative example of human problems with formal systems and formal reasoning is presented by the Wason selection task. It is one of

the cornerstones of evolutionary psychology. (See [3] for more information and references.) In this task, experimental subjects are shown four cards lying on a table. Each card is about a particular individual, Alice, Bob, Charlie, or Donna, and on on each side a statement about act by that individual. The subject's task is to decide, after reading the top sides of the cards, which of these cards need to be turned over to find out whether that individual satisfied some explicit rule. For example, we might be told that Alice, Bob, Charlie, and Donna all live in Philadelphia, and the rule might be that "If a person travels from Philadelphia to Chicago, he or she flies." For each person, one side of the card states where that person went, the other how they got there. The top side of Alice's card might say that she traveled to Baltimore, Bob's might say that he drove a car, Charlie's that he went to Chicago, and Donna's that she flew. For a logically minded person, it is clear that it is precisely Bob's and Charlie's cards that have to be turned over. In practice, though, only about a quarter of all subjects manage to figure this out.

The surprising part of the Wason selection task is what happens when the problem is restated. Suppose that now Alice, Bob, Charlie, and Donna are said to be children in a family, and the parents have a rule that "If a child has ice cream for dessert, he or she has to do the dishes after the meal." Suppose next that the top side of Alice's card states that she had fruit for dessert, Bob's that he watched TV after the meal, Charlie's that he had ice cream, and Donna's that she did the dishes. Most technologists immediately say that this is exactly the same problem as before, with only the wording changed. The rule is still of the form "If X then Y," only X and Y are different in the two cases, so again it is precisely Bob's and Charlie's cards that have to be turned over to check whether the rule is satisfied. Yet, among the general population, about three quarters manage to get this task right, in comparison to just one quarter for the earlier version. This (together with other experiments with other wordings and somewhat different settings) is interpreted as indicating that we have specialized mental circuits for detecting cheating in social settings.

The extended discussion of most people's difficulties with formal methods is motivated by the fact that security systems are conceived, developed, and deployed by technologists. They are among the small fraction of the human race that is comfortable with formal systems. They usually have little patience for human factors and social relations. In particular, they tend to expect others to think the way they do, and to be skilled at the formal thinking that the design and proper operation of secure systems require.

While people do have trouble with formal reasoning, we should not forget that they are extremely good at many tasks that computers are poor at. Just about any four year old girl is far superior in the ability to speak, understand spoken language, or recognize faces to even the most powerful and sophisticated computer system we have been able to build. Such abilities enable people to function in social settings, and in particular to cope with insecure systems. In particular, since information and communication systems do not operate in iso-

lation, and instead are at the service of a complicated society, there is a context to most electronic transactions that provides an extra margin of safety.

3 Digital Signatures versus Fax Signature

The 1980s were the golden age of civilian research on cryptography and security. The seeds planted in the 1970s were sprouting, and the technologists' bright hopes for a brave new world had not yet collided with the cold reality as clearly as they did in the 1990s. Yet the 1980s were also the age of the fax, which became ubiquitous. With the fax, we got fax signatures. While security researchers were developing public key infrastructures, and worrying about definitions of digital signatures, fax signatures became widespread, and are now playing a crucial role in the economy. Yet there is practically nothing as insecure as a fax signature, from a formal point of view. One can easily copy a signature from one document to another and this will be imperceptible on a fax. So what lessons can we draw from fax signatures, other than that convenience trumps security? One lesson is that the definition of a signature is, as with the message "Please let in the plumber to fix the water heater," loaded with cultural baggage that is hard to formalize.

It turns out that there is no strict legal definition of ordinary signature. We may think we know what a valid signature is, but the actual situation is quite complicated. An "X" may very well be a valid signature, even if it comes from somebody who normally signs her name in full. (She may have her hand in a cast, for example.) On the other hand, a very ordinary signature may not be valid, say if the signer was drunk while making it, or had a gun held to her head.

Furthermore, any signature, digital or physical, even if made willingly, may not be regarded as valid for legal enforcement of contract. Minors are not allowed to enter into most contracts. Even adults are not allowed to carry out some contracts that are regarded as against social policy, such as selling themselves or their children into slavery. Our legal system embodies many cultural norms (which vary from society to society, and even within a society change with time).

Another lesson is that our society somehow managed to function even when signatures became manifestly less secure with the spread of fax signatures. Moreover, it is easy to argue that fax signatures have contributed greatly to economic growth. How did this happen? This occurred because there is a context to almost every fax communication.

4 Social, Legal, and Economic Checks and Balances

Although fax signatures have become widespread, their usage is restricted. They are not used for final contracts of substantial value, such as home purchases. That means that the insecurity of fax communication is not easy to exploit for large gain. Additional protection against abuse of fax insecurity is provided by the context in which faxes are used. There are records of phone calls that carry the faxes, paper trails inside enterprises, and so on. Furthermore, unexpected large

financial transfers trigger scrutiny. As a result, successful frauds are not easy to carry out by purely technical means. Insiders (as at Enron and WorldCom and innumerable other enterprises) are much more dangerous.

Our commercial, government, and academic enterprises are large organizations with many formal rules and regulations. Yet the essential workings of these enterprises are typically based on various social relations and unwritten rules. As a result, one of the most effective tactics that employees have in pressuring management in labor disputes is to "work to rule." In general, the social and organizational aspects of large enterprises and even whole economies are poorly understood and underappreciated. Standard quantitative measures of invested capital or access to technology do not explain phenomena such as the continuing substantial lag of the regions of former East Germany behind former West Germany. There are other puzzling observations, such as the typical lack of measurable impact on economic output from major disruptions, such as earthquakes and snowstorms. There are ongoing attempts to understand just how societies function, including explorations of novel concepts such as "social capital." In general, though, it has to be said that our knowledge is still slight.

Information and communication technologies do play a crucial role in enabling smooth functioning of our complicated society, but are just a small part of it. That provides natural resilience in the face of formal system insecurities. Furthermore, the same limitations that make it hard to design, deploy, and effectively run secure systems also apply to attackers. Most criminals are stupid. Even those that are not stupid find it hard to observe the security precautions that are required for successful crime (such as inconspicuous consumption of their illicit gains). Even as determined an attacker as al Qaeda has had numerous security breaches. And, of course, the usual economic incentives apply to most attackers, namely that they are after material gains, have limited resources, and so on.

The natural resilience of human society suggests yet again the natural analogies between biological defense systems and technological ones. An immune system does not provide absolute protection in the face of constantly evolving adversaries, but it provides adequate defense most of the time.

The standard thinking in information security has been that absolute security is required. Yet we do have a rapidly growing collection of data that shows the value of even imperfect security. The experience of the pay-TV industry is certainly instructive. Although their systems have been cracked regularly, a combination of legal, technological, and business methods has kept the industry growing and profitable. Some more examples are offered by the applications of encryption technologies to provide lock-in for products, as in the replacement printer cartridge market [9]. Very often, "speed bumps" is all that is needed to realize economic value.

5 Conclusions

The general conclusion is that there is no "silver bullet" for security. In a society composed of people who are unsuited to formally secure systems, the best we

can hope to do is to provide "speed bumps" that will reduce the threat of cyberattacks to that we face from more traditional sources.

References

1. Anderson, R.J.: Liability and Computer Security - Nine Principles. ESORICS 94. Available at ⟨http://www.cl.cam.ac.uk/∼rja14⟩.
2. Anderson, R.J.: Security Engineering - A Guide to Building Dependable Distributed Systems. Wiley, 2001.
3. Cosmides, L., Tooby, J.: Evolutionary Psychology: A Primer. Available at ⟨www.psych.ucsb.edu/research/cep/primer.html⟩.
4. Geer, D.: Risk Management is Where the Money Is. Risks Digest, vol. 20, no. 6, Nov. 12, 1998. Available at ⟨http://catless.ncl.ac.uk/Risks/20.06.html⟩.
5. Norris, F.: No positives in this legal double negative. New York Times, January 24, 2003.
6. Odlyzko, A.M.: The Bumpy Road of Electronic Commerce. In: Maurer, H. (ed.): WebNet 96 - World Conf. Web Soc. Proc.. AACE (1996) 378–389. Available at ⟨http://www.dtc.umn.edu/∼odlyzko/doc/recent.html⟩.
7. Odlyzko, A.M.: The Visible Problems of the Invisible Computer: A Skeptical Look at Information Appliances. First Monday, **4** (no. 9) (Sept. 1999), ⟨http://www.firstmonday.org/issues/issue4_9/odlyzko/index.html⟩. Also available at ⟨http://www.dtc.umn.edu/∼odlyzko/doc/recent.html⟩.
8. Odlyzko, A.M.: Cryptographic Abundance and Pervasive Computing. iMP: Information Impacts Magazine, June 2000, ⟨http://www.cisp.org/imp/june_2000/06_00odlyzko-insight.htm⟩. Also available at ⟨http://www.dtc.umn.edu/∼odlyzko/doc/recent.html⟩.
9. Static Control Corporation: Computer Chip Usage in Toner Cartridges and Impact on the Market: Past, Current and Future. White paper, dated Oct. 23, 2002, available at ⟨http://www.scc-inc.com/special/oemwarfare/default.htm⟩.
10. Schneier, B.: Secrets and Lies: Digital Security in a Networked World. Wiley, 2000.
11. Workshop on Economics and Information Security: May 16-17, 2002. Program and papers or abstracts available at ⟨http://www.sims.berkeley.edu/resources/affiliates/workshops/econsecurity/⟩.