The Economics of Information Security

Ross Anderson and Tyler Moore

University of Cambridge, Computer Laboratory
15 JJ Thomson Avenue, Cambridge CB3 0FD, United Kingdom
firstname.lastname@cl.cam.ac.uk

The economics of information security has recently become a thriving and fast-moving discipline. As distributed systems are assembled from machines belonging to principals with divergent interests, we find incentives becoming as important to dependability as technical design is. The new field provides valuable insights not just into 'security' topics such as bugs, spam, phishing and law-enforcement strategy, but into more general areas such as the design of peer-to-peer systems, the optimal balance of effort by programmers and testers, why privacy gets eroded, and the politics of DRM.

Introduction

Over the last six years, people have realized that security failure is caused at least as often by bad incentives as by bad design. Systems are particularly prone to failure when the person guarding them is not the person who suffers when they fail. The growing use of security mechanisms to enable one system user to exert power over another user, rather than simply to exclude people who should not be users at all, introduces many strategic and policy issues. The tools and concepts of game theory and microeconomic theory are becoming just as important to the security engineer as the mathematics of cryptography.

We review several recent results and live research challenges in the economics of information security. We first consider misaligned incentives, and externalities: network insecurity is somewhat like air pollution or traffic congestion, in that people who connect insecure machines to the Internet do not bear the full consequences of their actions.

The difficulty in measuring information security risks presents another challenge: these risks cannot be managed better until they can be measured better. Auctions and markets can help reduce the information asymmetry prevalent in the software industry. We also examine the problem of insuring against attacks. The local and global correlations exhibited by different attack types largely determine what sort of insurance markets are feasible. Information security

mechanisms or failures can create, destroy or distort other markets: DRM provides a topical example.

Economic factors also explain many challenges to personal privacy. Discriminatory pricing – which is economically efficient but socially controversial – is simultaneously made more attractive to merchants, and easier to implement, by technological advance. We conclude by discussing a fledgling research effort: examining the security impact of network structure on interactions, reliability and robustness.

Misaligned incentives

One of the observations that drove initial interest in security economics came from banking. In the USA, banks are generally liable for the costs of card fraud; when a customer disputes a transaction, the bank must either show that she is trying to cheat, or refund her money. In the UK, the banks had a much easier ride: they generally got away with claiming that the ATM system was 'secure', so a customer who complained must be mistaken or lying. "Lucky bankers," you might think; yet UK banks spent more on security and suffered more fraud. How could this be? It appears to have been what economists call a moral-hazard effect: UK bank staff knew that customer complaints would not be taken seriously, so they became lazy and careless. This situation led to an avalanche of fraud [1].

In 2000, Varian made another key observation – about the anti-virus software market. People did not spend as much on protecting their computers as they might have. Why not? Well, at that time, a typical virus payload was a service-denial attack against the website of a company like Microsoft. While a rational consumer might well spend $20 to prevent a virus from trashing his hard disk, he might not do so just to prevent an attack on someone else [2].

Legal theorists have long known that liability should be assigned to the party that can best manage the risk. Yet everywhere we look, we see online risks allocated poorly, resulting in privacy failures and protracted regulatory tussles. For instance, medical IT systems are bought by hospital directors and insurance companies, whose interests in account management, cost control and research are not well aligned with the patients' interests in privacy.

Incentives can also influence attack and defense strategies. In economic theory, a hidden-action problem arises when two parties wish to transact, but one party can take unobservable actions that impact the transaction. The classic example comes from insurance, where the insured party may behave recklessly (increasing the likelihood of a claim) because the insurance company cannot observe his behavior.

Moore noted that we can use such economic concepts to classify computer security problems [3]. Routers can quietly drop selected packets or falsify responses to routing requests; nodes can redirect network traffic to eavesdrop on conversations; and players in file-sharing systems can hide whether they have chosen to share with others, so some may 'free-ride' rather than to help sustain the system. In such *h*idden-action attacks, some nodes can hide malicious or antisocial behavior from others. Once the problem is seen in this light, designers can structure

interactions to minimize the capacity for hidden action or to make it easy to enforce suitable contracts.

This helps explain the evolution of peer-to-peer systems over the past ten years. Early systems, such as Eternity, Freenet, Chord, Pastry and OceanStore, provided a 'single pot', with widely and randomly distributed functionality. Later and more successful systems, like the popular Gnutella and Kazaa, allow peer nodes to serve content they have downloaded for their personal use, without burdening them with random files. The comparison between these architectures originally focused on purely technical aspects: the cost of search, retrieval, communications and storage. However, it turns out that incentives matter here too.

First, a system structured as an association of clubs reduces the potential for hidden action; club members are more likely to be able to assess correctly which members are contributing. Second, clubs might have quite divergent interests. Early peer-to-peer systems were oriented towards censorship resistance rather than music file sharing, and when they put all content into one pot, quite different groups ended up protecting each others' free speech – maybe Chinese dissidents, critics of Scientology, or aficionados of sado-masochistic imagery that is legal in California but banned in Tennessee. The question then is whether such groups might not fight harder to defend their own kind, rather than people involved in struggles in which they had no interest and where they might even be disposed to side with the censor.

Danezis and Anderson introduced the Red-Blue model to analyze this [4]. Each node has a preference among resource types, while a censor who attacks the network will try to impose his own preferences. His action will meet the approval of some nodes but not others. The model proceeds as a multi-round game in which nodes set defense budgets which affect the probability that they will defeat the censor or be overwhelmed by him. Under reasonable assumptions, the authors show that diversity (with each node storing its preferred resource mix) performs better under attack than solidarity (where each node stores the same resource mix, which is not usually its preference). Diversity increases node utility which in turn makes nodes willing to allocate higher defense budgets. This model sheds light on the more general problem of the tradeoffs between diversity and solidarity when conflict threatens, and the related social policy issue of the extent to which the growing diversity of modern societies is in tension with the solidarity on which modern welfare systems are founded [5].

Security as an externality

Information industries are characterized by many different types of *externalities*, where individuals' actions have side-effects on others. The software industry tends toward dominant firms thanks to the benefits of interoperability. Economists call this a network externality: a network, or a community of software users, is more valuable to its members the larger it is. This not only helps explain the rise and dominance of operating systems, from System/360 through Windows to Symbian, and of music platforms such as iTunes; it also helps explain the typical pattern of security flaws. Put simply, while a platform vendor is building market dominance, he has to

appeal to vendors of complementary products as well as to his direct customers; not only does this divert energy that might be spent on securing the platform, but security could get in the way by making life harder for the complementers. So platform vendors commonly ignore security in the beginning, as they are building their market position; later, once they have captured a lucrative market, they add excessive security in order to lock their customers in tightly [7].

Another instance of externalities can be found when we analyze security investment, as protection often depends on the efforts of many principals. Budgets generally depend on the manner in which individual investment translates to outcomes, but this in turn depends not just on the investor's own decisions but also on the decisions of others.

System reliability can depend on the sum of individual efforts, the minimum effort anyone makes, or the maximum effort any makes. Program correctness can depend on the weakest link (the most careless programmer introducing a vulnerability) while software validation and vulnerability testing might depend on the sum of everyone's efforts. There can also be cases where the security depends on the best effort – the effort of an individual champion. A simple model by Varian provides interesting results when players choose their effort levels independently [8]. For the total-effort case, system reliability depends on the agent with the highest benefit-cost ratio, and all other agents free-ride. In the weakest-link case, the agent with the lowest benefit-cost ratio dominates. As more agents are added, systems become increasingly reliable in the total-effort case but increasingly unreliable in the weakest-link case. What are the implications? One is that software companies should hire more software testers and fewer (but more competent) programmers.

Work such as this has inspired other researchers to consider interdependent risk. A recent influential model by Kunreuther and Heal notes that the security investments can be strategic complements: an individual taking protective measures creates positive externalities for others that in turn may discourage their own investment [9]. This result has implications far beyond information security. The decision by one apartment owner to install a sprinkler system that minimizes the risk of fire damage will affect the decisions of his neighbors; airlines may decide not to screen luggage transferred from other carriers who are believed to be careful with security; and people thinking of vaccinating their children against a contagious disease may choose to free-ride off the herd immunity instead. In each case, several widely varying Nash equilibrium outcomes are possible, from complete adoption to total refusal, depending on the levels of coordination between principals.

Katz and Shapiro famously noted how network externalities affected the adoption of technology [10]. Network effects can also influence the deployment of security technology. The benefit a protection technology provides may depend on the number of users that adopt it. The cost may be greater than the benefit until a minimum number of players adopt; so each decision-maker might wait for others to go first, and the technology never gets deployed. Recently, Ozment and Schechter have analyzed different approaches for overcoming bootstrapping problems faced by those who would deploy security technologies [11].

This challenge is particularly topical. A number of core Internet protocols, such as DNS and routing, are considered insecure. More secure protocols exist; the challenge is to get them

adopted. Two security protocols that have already been widely deployed, SSH and IPsec, both overcame the bootstrapping problem by providing significant intra-organizational benefits. In the successful cases, adoption could be done one organization at a time, rather than needing most organizations to move at once. The deployment of fax machines also occurred through this mechanism: companies initially bought fax machines to connect their own offices.

Economics of vulnerabilities

A vigorous debate has ensued between software vendors and security researchers over whether actively seeking and disclosing vulnerabilities is socially desirable. Resorla has argued that for software with many latent vulnerabilities (like Windows), removing an individual bug makes little difference to the likelihood of an attacker finding another one later [12]. Since exploits are often based on vulnerabilities inferred from patches or security advisories, he argued against disclosure and frequent patching if vulnerabilities are correlated.

Ozment investigated vulnerabilities identified for FreeBSD; he found that many vulnerabilities are indeed likely to be rediscovered and are therefore often correlated [13]. Arora, Telang and Xu produced a model where disclosure is necessary to incentivize vendors into fixing bugs in subsequent product releases [14]. Arora, Krishnan, Nandkumar, Telang and Yang present quantitative analysis to complement the above model, which found that for public disclosure, vendors respond more quickly compared to private disclosure, the number of attacks increases but the number of reported vulnerabilities declines over time [15].

This discussion begs a more fundamental question: why do so many vulnerabilities exist in the first place? Surely, if companies desire secure products then secure software will dominate the marketplace? As we know from experience, this is not the case: most commercial software contains design and implementation flaws that could easily have been prevented. Although vendors are capable of creating more secure software, the economics of the software industry provide them with little incentive to do so [7]. In many markets, the attitude of 'ship it Tuesday and get it right by version 3' is perfectly rational behavior. Consumers generally reward vendors for adding features, for being first to market, or for being dominant in an existing market – and especially so in platform markets with network externalities. These motivations clash with the task of writing more secure software, which requires time-consuming testing and a focus on simplicity.

Another aspect of vendors' lack of motivation is readily explained by Anderson: the software market is a 'market for lemons.' In a Nobel prize-winning work, economist George Akerlof employed the used car market as a metaphor for a market with asymmetric information [16]. His paper imagines a town in which 50 good used cars (worth $2000) are for sale, along with 50 'lemons' (worth $1000) each). The sellers know the difference but the buyers do not. What will be the market-clearing price? One might initially think $1500, but at that price no-one with a good car will offer it for sale; so the market price will quickly end up near $1000. Because buyers are unwilling to pay a premium for quality they cannot measure, only

low-quality used vehicles are available for sale.

The software market suffers from the same information asymmetry. Vendors may make claims about the security of their products, but buyers have no reason to trust them. In many cases, even the vendor does not know how secure its software is. So buyers have no reason to pay more for more secure software, and vendors are disinclined to invest in protection. How can this be tackled?

There are two developing approaches to obtaining accurate measures of software security: vulnerability markets and insurance.

Vulnerability markets help buyers and sellers to establish the actual cost of finding a vulnerability in software, which is a reasonable proxy for software security. To begin with, some standards specified a minimum cost of various kinds of technical compromise; one example is banking standards for PIN-entry terminals [17]. Then Schechter proposed open markets for reports of previously undiscovered vulnerabilities [18]. Two organizations are now openly buying vulnerabilities, so a market actually exists (unfortunately, the prices are not published). Their business model is to provide vulnerability data simultaneously to their customers and to the vendor of the affected product, so that their customers can update their firewalls before anyone else. Kannan and Telang have analyzed the social utility of this and found it to be suboptimal [20]: bug-market organizations may have an incentive to leak vulnerability information without proper safeguards.

Böhme has argued that software derivatives are a better tool than markets for the measurement of software security [21]. Here, security professionals can reach a price consensus on the level of security for a product. Contracts for software could be issued in pairs: the first pays a fixed value if no vulnerability is found in a program by a specific date, and the second pays another value if vulnerabilities are found. If these contracts can be traded, then their price will reflect the consensus on the program. Software vendors, software company investors, and insurance companies could use such derivatives to hedge risks. A third possibility, due to Ozment, is to design a vulnerability market as an auction [19].

One criticism of market-based approaches is that they might increase the number of identified vulnerabilities by compensating people who would otherwise not search for flaws. Thus, some care must be exercised in designing them.

An alternative approach is to rely on insurers. The argument is that underwriters assign premiums based upon a firm's IT infrastructure and the processes by which it is managed. This assessment results in both detailed best practices and, over the long run, a pool of data by which the insurer can value risks accurately. Right now, however, the cyber-insurance market is both underdeveloped and underutilized. Why could this be?

One reason, according to Böhme and Kataria [22], is the problem of interdependent risk, which takes at least two forms. Firms' IT infrastructure is connected to other entities – so its efforts may be undermined by failures elsewhere. Cyber attacks also often exploit a vulnerability in a system used by many firms. This interdependence makes certain cyber-risks unattractive to insurers – particularly those where the risk is globally rather than locally correlated, such as worm and virus attacks, and systemic risks such as Y2K. Now many writers have called for soft-

ware risks to be transferred to the vendors; but if this were the law, it is unlikely that Microsoft would be able to buy insurance. So far, vendors have succeeded in dumping most software risks; but this outcome is also far from being socially optimal. Even at the level of customer firms, correlated risk makes firms underinvest in both security technology and cyber-insurance [23]. nsurance companies must charge higher premiums, and so cyber-insurance markets lack the volume and liquidity to become efficient.

Insurance is not the only market affected by information security. Some very high-profile debates have centred on DRM; record companies have pushed for years to be incorporated into computers and consumer electronics, while digital-rights activists have opposed them. What light can security economics shed on this debate?

Varian presented a surprising result in January 2005 [6]: that stronger DRM would help platform cendors more than the music industry, because the computer industry is more concentrated (with only three serious DRM suppliers – Microsoft, Sony, and the dominant firm, Apple). The content industry scoffed, but by the end of the year music publishers were protesting that Apple was getting an unreasonably large share of the cash from online music sales. As power in the supply chain moved from the music majors to the platform vendors, so power in the music industry appears to be shifting from the majors to the independents, just as airline deregulation has favoured aircraft makers and low-cost airlines. This is a striking demonstration of the predictive power of economic analysis.

There are other interesting market failures. Recently, for example, a number of organizations have set up certification services to vouch for the quality of software products or web sites. The aim has been twofold: to overcome public wariness about electronic commerce, and by self-regulation to forestall more expensive regulation by the government. But certification markets can easily be ruined by a race to the bottom; dubious companies are more likely to purchase certificates than reputable ones, and even ordinary companies may shop around for the easiest deal. Edelman has shown that this such 'adverse selection' is really happening [24]: while about 3% of websites are malicious, some 8% of websites with certification from one large vendor are malicious. He also discovered inconsistencies between ordinary web search results and those from paid advertising, finding that while 2.73% of companies ranked top in a web search were bad, 4.44% of companies who had bought ads from the search engine were bad. His conclusion: 'Don't click on ads'.

Economics of privacy

The persistent erosion of personal privacy with advances in technology has frustrated policy people and practitioners alike. Privacy-enhancing technologies have been offered for sale, yet most have failed in the marketplace. Again, economics explains this better than technical factors.

Odlyzko argues that privacy erosion is a consequence of the desire to charge different prices for similar services [25]. Technology is increasing both the incentives and the opportunities

for discriminatory pricing. Companies can mine online purchases and interactions for data revealing individuals' willingness to pay. From airline yield management systems to complex and ever-changing software and telecommunications prices, differential pricing is economically efficient – but increasingly resented. Acquisti and Varian analyzed the market conditions under which first-degree price discrimination can actually be profitable [26]: it may thrive in industries with wide variation in consumer valuation for services, where personalized services can be supplied with low marginal costs, and where repeated purchases are likely.

So much for the factors that make privacy intrusions more likely. What factors make them less so? Campbell, Gordon, Loeb and Zhou found that the stock price of companies reporting a security breach is more likely to fall if the breach leaked confidential information [27]. Acquisti, Friedman and Telang conducted a similar analysis for privacy breaches [28]. Their initial results are less conclusive but still point to a negative impact on stock price followed by an eventual recovery.

Incentives also affect the detailed design of privacy technology. Anonymity systems depend heavily on network externalities: additional users provide cover traffic necessary to hide users' activities from an observer. This fact has been recognized by some developers of anonymity systems [29]. As a result, some successful applications like Tor [30], which anonymizes web traffic, emphasize usability to increase adoption rates.

On the horizon: network topology and information security

The topology of complex networks is an emerging tool for analyzing information security. Computer networks from the Internet to decentralized peer-to-peer networks are complex but emerge from ad-hoc interactions of many entities using simple ground rules. This emergent complexity, coupled with heterogeneity, is similar to social networks made up from interactions between people, and even the metabolic pathways in living organisms. Recently a discipline of network analysis has emerged at the boundary between sociology and condensed-matter physics. It takes ideas from other disciplines like graph theory, and in turn provides tools for modeling and investigating such networks (see [31] for a recent survey). The interaction of network science with information security provides an interesting bridge to evolutionary game theory, a branch of economics that has been very influential in the study of human and animal behaviour.

Network topology can strongly influence conflict dynamics. Often an attacker tries to disconnect a network or increase its diameter by destroying nodes or edges, while the defender counters using various resilience mechanisms. Examples include a music industry body attempting to close down a peer-to-peer file-sharing network; a police force trying to decapitate a terrorist organization; and a totalitarian government conducting surveillance on political activists. Police forces have been curious for some years about whether network science might be of practical use in covert conflicts – whether to insurgents or to counterinsurgency forces.

Different topologies have different robustness properties with respect to various attacks. Albert, Jeong and Barabási showed that certain real world networks with scale-free degree

distributions are more robust to random attacks than targeted attacks [32]. This is because scale-free networks – like many real-world networks – get much of their connectivity from a minority of nodes that have a high vertex order. This resilience makes them highly robust against random upsets; but remove the 'kingpin' nodes, and connectivity collapses.

This is the static case – for example, when a police force becomes aware of a criminal or terrorist network, and sets out to disrupt it by finding and arresting its key people. Nagaraja and Anderson extend this to the dynamic case. In their model, the attacker can remove a certain number of nodes at each round, after which the defenders recruit other nodes to replace them [33]. They studied how attack and defense interact using multi-round simulations, and found that forming localized clique structures at key network points works reasonably well while defenses based on rings did not work well at all. This helps explain why peer-to-peer systems with ring architectures turned out to be rather fragile – and also why revolutionaries have tended to organize themselves in cells.

Conclusion

Over the last few years, a research program on the economics of security has built many cross-disciplinary links and has produced many useful (and indeed delightful) insights from unexpected places.

We have discussed how many information security failures are caused by incentive failures, for example where the people who guard a system aren't the people who suffer when it fails. Externalities make many security problems somewhat like environmental pollution; some aspects of information security are public goods, like clean air and water. Externalities also play a key role in determining which security products succeed in the market, and which fail.

Games with incomplete information also play an important role: where either information or action is hidden, things can go wrong in interesting ways. Markets, and auctions, can sometimes be used as information processing mechanisms to tackle the resulting problems; we discussed software dependability and the problems of cyber-insurance.

Economic explanations for privacy erosion also resonate, as increasing capacity for differential pricing is made possible by advances in information technology. Finally, the global structure resulting from localized interactions is often susceptible to manipulation. This is an exciting new opportunity for all kinds of research.

References

[1] R. J. Anderson, *Communications of the ACM* **37**, 32 (1994).

[2] H. Varian, *The New York Times* (2000). `http://www.nytimes.com/library/financial/columns/060100econ-scene.html`.

[3] T. Moore, *Proceedings of the Fourth Workshop on the Economics of Information Security* (2005).

[4] G. Danezis, R. J. Anderson, *IEEE Security & Privacy* **3**, 45 (2005).

[5] D. Goodhart, *Prospect* (2004). http://www.guardian.co.uk/race/story/0,11374,1154684,00.html.

[6] H. Varian. Keynote address to the Third Digital Rights Management Conference, Berlin, Germany, January 13, 2005.

[7] R. Anderson, *17th Annual Computer Security Applications Conference* (2001).

[8] H. Varian, *Economics of Information Security*, L. J. Camp, S. Lewis, eds. (Kluwer Academic Publishers, 2004), vol. 12 of *Advances in Information Security*, pp. 1–15.

[9] H. Kunreuther, G. Heal, *Journal of Risk and Uncertainty* **26**, 231 (2003).

[10] M. L. Katz, C. Shapiro, *The American Economic Review* **75**, 424 (1985).

[11] A. Ozment, S. E. Schechter, *Proceedings of the Fifth Workshop on the Economics of Information Security* (2006).

[12] E. Resorla, *Proceedings of the Third Workshop on the Economics of Information Security* (2004).

[13] A. Ozment, *Proceedings of the Fourth Workshop on the Economics of Information Security* (2005).

[14] A. Arora, R. Telang, H. Xu, *Proceedings of the Third Workshop on the Economics of Information Security* (2004).

[15] A. Arora, R. Krishnan, A. Nandkumar, R. Telang, Y. Yang, *Proceedings of the Third Workshop on the Economics of Information Security* (2004).

[16] G. A. Akerlof, *The Quarterly Journal of Economics* **84**, 488 (1970).

[17] PIN management requirements: PIN entry device security requirements manual (2004). http://partnernetwork.visa.com/dv/pin/pdf/Visa_ATM_Security_Requirements.pdf.

[18] S. Schechter, Computer security strength & risk: A quantitative approach, Ph.D. thesis, Harvard University (2004).

[19] A. Ozment, *Proceedings of the Third Workshop on the Economics of Information Security* (2004).

[20] K. Kannan, R. Telang, *Proceedings of the Third Workshop on the Economics of Information Security* (2004).

[21] R. Böhme, *Proceedings of ETRICS* (Springer Verlag, 2006), pp. 298–311. LNCS 2995.

[22] R. Böhme, G. Kataria, *Proceedings of the Fifth Workshop on the Economics of Information Security* (2006).

[23] H. Ogut, N. Menon, S. Raghunathan, *Proceedings of the Fourth Workshop on the Economics of Information Security* (2005).

[24] B. Edelman, *Proceedings of the Fifth Workshop on the Economics of Information Security* (2006).

[25] A. Odlyzko, *ICEC '03: Proceedings of the 5th international conference on Electronic commerce* (ACM Press, New York, NY, USA, 2003), pp. 355–366.

[26] A. Acquisti, H. Varian, *Marketing Science* **24** (2005).

[27] K. Campbell, L. A. Gordon, M. P. Loeb, L. Zhou, *J. Comput. Secur.* **11**, 431 (2003).

[28] A. Acquisti, A. Friedman, R. Telang, *Proceedings of the Fifth Workshop on the Economics of Information Security* (2006).

[29] R. Dingledine, N. Matthewson, *Workshop on Usable Privacy and Security Software* (2004).

[30] `http://tor.eff.org`.

[31] M. E. J. Newman, *SIAM Review* **45**, 167 (2003).

[32] R. Albert, H. Jeong, A. lászló Barabási, *Nature* **406**, 387 (2000).

[33] S. Nagaraja, R. Anderson, *Proceedings of the Fifth Workshop on the Economics of Information Security* (2006).

[34] L. Li, D. Alderson, W. Willinger, J. Doyle, *SIGCOMM*, R. Yavatkar, E. W. Zegura, J. Rexford, eds. (ACM, 2004), pp. 3–14.