# The Iterated Weakest Link

Rainer Böhme [a] and Tyler Moore [b]

[a] Technische Universität Dresden, Inst. of Systems Architecture, Germany;
rainer.boehme@tu-dresden.de

[b] Harvard University, Center for Research on Computation and Society, USA;
tmoore@seas.harvard.edu

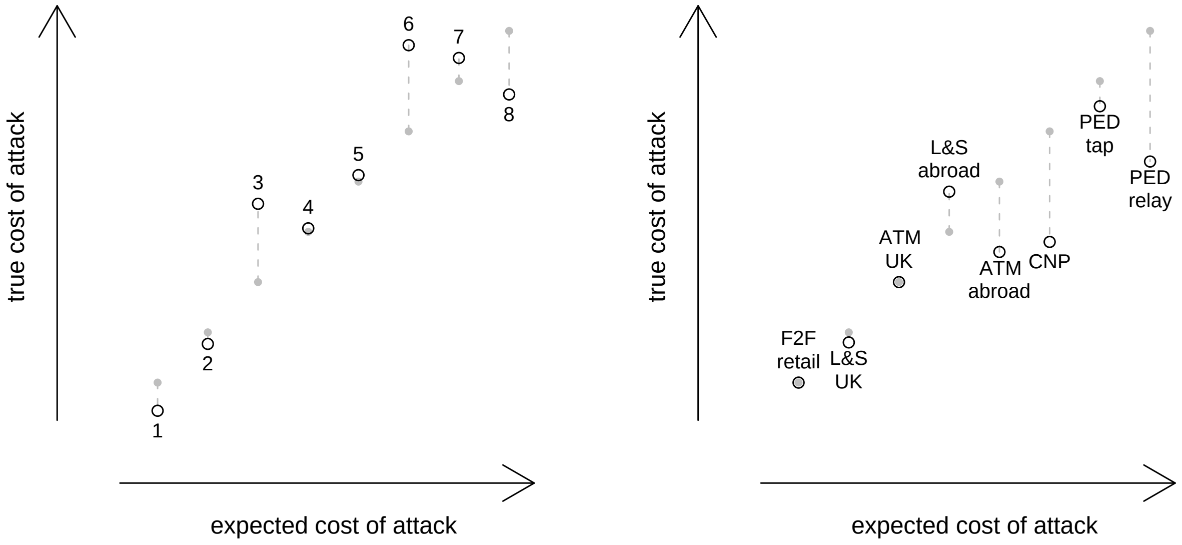
Security breaches are in the news almost daily, each bigger and more costly than the last. Does this reflect flawed technology, policy, or simply ineptitude? What if, instead, allowing some attacks to succeed is entirely rational? Rather than over-invest proactively, companies could wait to observe which attacks work and use this knowledge to better allocate security spending. In this essay, we describe a model that weighs the merits of such an approach.[1]

One key insight from the economics of information security literature [2] is that attackers bent on undermining a system's security operate *strategically.* Moreover, information systems are often structured so that a system's overall security depends on its *weakest link* [3]. The most careless programmer in a software firm can introduce a critical vulnerability. The Internet's global, distributed architecture leads to security being dominated by the weakest link. Attackers have repeatedly exhibited a knack for identifying ways to bypass a system's security, even when the system's designer remains unaware of the particular weakness.

However, systems do not exist in a vacuum; rather, defenders respond to attacks by plugging known holes. And yet, as soon as one flaw is fixed, another weak point is often identified and exploited. Therefore, a strong dynamic component is at play: attackers find the weakest link, defenders fix the problem, attackers find new holes which are then plugged, and so on. We see this pattern emerge repeatedly. For instance, attackers construct networks of compromised machines (so-called botnets) to pester legitimate users by emitting spam, distributing malware and hosting phishing websites. Attackers concentrate their efforts at the most irresponsible ISPs, moving on to others only after the ISP cleans up its act or is shut down [4, 5]. Likewise, technical countermeasures to payment card fraud have evolved over time, causing fraudsters to adopt new strategies as old weaknesses are fixed. For example, when UK banks migrated to PIN verification of transactions rather than signatures, in-person retail fraud declined while overseas ATM fraud and card-not-present fraud skyrocketed [6].

So how can we grasp and model this dynamic interaction between attackers and defenders? Simply stated, a defender protects an asset of value against $n$ possible threats. Each

[1] This essay was selected as winner of the inaugural Gordon Prize in Managing Cybersecurity Resources (http://www.rhsmith.umd.edu/news/releases/2009/101409.aspx). A full academic paper outlining the model presented here can be found in [1].



(a) Gaussian uncertainty: $\sigma = 1$

(b) EMV attack costs

Figure 1: Example of uncertainty for attack costs (hypothetical values)

threat can be warded off by investing in its corresponding defense. The costs for each defensive countermeasure are represented by an upper triangular matrix. Costs can be modeled as independent (off-diagonal values zero), complementary (off-diagonal values negative), or conflicting (off-diagonal values positive). One nice property of arranging the cost matrix in this manner is that for positive off-diagonal elements, decreasing marginal utility of defenses becomes endogenous instead of appearing as an assumption as in the Gordon–Loeb framework (A3, p. 443 in [7]).

Whether interdependent or not, the costs to the defender of implementing countermeasures are assumed to be known. This is reasonable – security countermeasures such as firewalls and intrusion detection systems come with a bill. By contrast, it is much harder for a defender to accurately predict the cost of different attacks in advance. While the defender may possess some intuition about the relative difficulty of carrying out the $n$ threats, such knowledge may very well be blurred.

To model this uncertainty, we order the threats $1, \ldots, n$ by increasing *expected* cost of attack. By varying the level of uncertainty $\sigma$ associated with the *true* costs for different attacks, we can learn a great deal about why security investment often falls short of what technical experts desire. Figure 1 (a) illustrates the role of uncertainty when ordering threats. Under uncertainty, expected and realized costs differ so that threat 4, not threat 3 as expected, is the weakest link if defenses 1 and 2 are in place.

To connect the model to a concrete example, consider the many threats to payment card security as presented in Figure 1 (b). Face-to-face retail fraud (F2F) might reasonably be seen as the weakest link in the payment card environment; its reduction following the adoption of

Chip and PIN supports this view. Similarly, the banks correctly anticipated that losses due to credit cards lost or stolen (L&S) inside the UK would drop once PINs were required for use. One area where the banks' expectations were not met is with ATM fraud on UK cards outside the UK. It turns out that fraudsters can easily clone stolen UK cards and use them in foreign ATMs; hence the true cost of attack is lower than expected. Likewise, the banks' losses due to card-not-present fraud (CNP) were much higher than forecast; unsurprisingly, many banks have now decided to deploy readers that verify PINs.

Returning to our model, it is 'run' in an iterated game; in each round, the defender decides which, if any, of the $n$ threats to protect against. The attacker identifies and exploits the weakest link, i. e., the threat least costly to the attacker. Unlike the defender, the attacker is certain of the cost for each attack. The attacker does not operate indiscriminately; rather, he only attacks when it is profitable to do so.

We reach several interesting conclusions upon examining the model. In the *static case*, where the defender only gets one chance to protect a system, increasing uncertainty about which link is weakest causes the defender to protect more assets, but only up to a point. When uncertainty is too high, the defender does not know which asset to protect and so chooses to protect none. If instead we allow for repeated defensive investments in the *dynamic case*, an uncertain defender will initially protect fewer assets and wait for the attacker to 'identify' the weakest links to be fixed in later rounds. Hence, it can be quite rational to *under-invest* in security until threats are realized.

Of course, security countermeasures may require significant capital investment from the outset. When we introduce sunk costs to our model [1], we find that for moderate levels of uncertainty, sunk costs raise the proactive protection investment adopted in the dynamic case.

We have translated our findings about optimal defensive strategies into accepted security indicators such as return on security investment (ROSI), as shown in Table 1. For moderate levels of uncertainty ($\sigma = 1$), moving from a static to dynamic defense strategy reduces security spending, which leads to more observed attacks. However, gross returns increase, too. So in fact, security spending is better targeted, over-investment is reduced, and the overall efficiency of security investment, as measured by the ROSI indicator, improves. Hence, we can draw an alternative interpretation to the omnipresent reports of security breaches in the media: rather than rashly framing them as engineering failures, one might also view breaches as unavoidable side-effects of smart defense strategies that balance the appropriate levels of proactive and reactive security investment.

Investment in countermeasures, and consequently the frequency of attack, depend fundamentally on the opportunity to defend reactively. When all security investment must be done proactively, firms may simply raise the white flag of surrender if they are very uncertain ($\sigma \geq 4$) about which threats are likely. Only a staged approach gives these investors an incentive to defend against the most aggressive threats. Given the chance to invest in later rounds, firms choose to protect the assets that have been revealed to be weak, leading to a higher return on investment and reduced attack intensity.[2]

[2] Our model identifies the rational response to the *private* costs faced by defenders, while ignoring the *public* costs created by insecurity. Hence, while it may be narrowly better for some defenders to skimp on security

| Indicator | Level of uncertainty<br>$\sigma = 0$ | <br>$\sigma = 1$ | <br>$\sigma = 4$ | <br>$\sigma = 8$ |
|---|---|---|---|---|
| **Static defense** | | | | |
| optimal number of defenses | 11 | 12 | 0 | 0 |
| attack intensity (% rounds) | 0.0 | 2.4 | 100.0 | 100.0 |
| ROSI (% security spending) | 51.5 | 31.2 | — | — |
| **Dynamic defense** | | | | |
| optimal number of proactive defenses | 11 | 9 | 7 | 3 |
| attack intensity (% rounds) | 0.0 | 6.1 | 15.7 | 32.7 |
| ROSI (% security spending) | 51.5 | 52.8 | 35.2 | 18.9 |

Table 1: Security investment indicators derived from model for sample parameters (asset value \$1 mil., return on asset 5%, loss given attack \$25 000, $n$=25, min. expected cost of attack \$15 000, gradient of attack cost \$1 000)

Our proposed economic model explains why and under which conditions security under-investment can be rational, even against known threats for which defenses exist. Unlike in other work explaining security under-investment with externalities or misaligned incentives, our model solely draws on uncertainty about *where* to invest in countermeasures. This result does not contradict or invalidate the well-known explanations by market failure [2]. It rather complements the picture and highlights that market failure is a sufficient, but not a necessary cause for security under-investment. The logic of initial security under-investment followed by reactive investment can also be found in real option frameworks that suggest a “wait-and-see” approach [8].

To conclude, we believe an iterated weakest link model accurately captures the challenges facing many information security threats today. Our findings suggest a need to reassess conclusions which condemn seemingly lax security practices observable in practice. Our model can assist policy makers in reducing negative externalities as consequences (not causes) of insecurity by better predicting situations where proactive investment is hindered. The model also helps identify influential factors, notably uncertainty about attacks, so that incentive-based countermeasures might be derived.

initially, a public policy response may nonetheless be necessary to compensate for the negative externalities of insecurity caused by such under-investment.

# References

[1] Böhme, R., Moore, T.: The iterated weakest link: A model of adaptive security investment. *Workshop on the Economics of Information Security (WEIS)*, University College London, UK (2009) `http://weis09.infosecon.net/files/152/paper152.pdf`.

[2] Anderson, R.J., Moore, T.: The economics of information security. *Science* **314** (2006) 610–613

[3] Varian, H.R.: System reliability and free riding. In Camp, L.J., Lewis, S., eds.: *Economics of Information Security*, Springer Verlag (2004) 1–15

[4] Moore, T., Clayton, R.: Examining the impact of website take-down on phishing. In: *Proc. of the Anti-Phishing Working Group eCrime Researchers Summit.* (2007) 1–13 `http://www.cl.cam.ac.uk/~rnc1/ecrime07.pdf`.

[5] Day, O., Palmen, B., Greenstadt, R.: Reinterpreting the disclosure debate for web infections. In Johnson, M.E., ed.: *Managing Information Risk and the Economics of Security*, New York, Springer (2008) 179–197

[6] APACS: 2007 UK Chip and PIN report (2007) `http://cryptome.org/UK-Chip-PIN-07.pdf`.

[7] Gordon, L.A., Loeb, M.P.: The economics of information security investment. *ACM Trans. on Information and System Security* **5** (2002) 438–457

[8] Gordon, L.A., Loeb, M.P., Lucyshyn, W.: Information security expenditures and real options: A wait-and-see approach. *Computer Security Journal* **14** (2003) 1–7