# The Plight of the Targeted Attacker in a World of Scale

Cormac Herley
Microsoft Research
One Microsoft Way
Redmond, WA, USA
cormac@microsoft.com

## ABSTRACT

Despite neglecting even basic security measures, close to two billion people use the Internet, and only a small fraction appear to be victimized each year. This paper suggests that an explanation lies in the economics of attacks. We distinguish between scalable attacks, where costs are almost independent of the number of users attacked, and non-scalable (or targeted) attacks, which involve per-user effort. Scalable attacks reach orders of magnitude more users. To compensate for her disadvantage in terms of reach the targeted attacker must target users with higher than average value.

To accomplish this she needs that value be both visible and very concentrated, with few users having very high value while most have little. In this she is fortunate: power-law longtail distributions that describe the distributions of wealth, fame and other phenomena are extremely concentrated. However, in these distributions only a tiny fraction of the population have above average value. For example, fewer than 2% of people have above average wealth in the US. Thus, when attacking assets where value is concentrated, the targeted attacker ignores the vast majority of users, since attacking them hurts rather than helps her requirement to extract greater than average value.

This helps explain why many users escape harm, even when they neglect security precautions: most users never experience most attacks. Attacks that involve per-user effort will be seen by only a tiny fraction of users. No matter how clever the exploit, unless the expected value is high, there is little place for per-user effort in this world of mass-produced attacks.

## 1. INTRODUCTION

An Internet user, Alice, must protect her resources from an attacker Charles. A common threat model assumes that Alice's strategy is known to Charles, who adapts in response to any changes she makes, or countermeasures she adopts. Thus, Alice's security is usually regarded as being only as strong as the weakest link. She must guard against every possible attack and patch every possible hole. Charles has many attacks that exploit vulnerabilities in her applications or operating system [44, 33, 10], her firewall [16], the network she uses [38] or her susceptibility to social engineering [27]. He even has techniques to spy on her using reflections from her LCD screen [34] or audio or electromagnetic emanations [29, 31]. He constantly adds new attacks. Unlike enterprizes which may be able to make weakest-link investment decisions [39, 30, 22] Alice may not even have approximate estimates of attack costs. Further, she may not know when a breach occurs (*e.g.*, if her PC is part of a botnet) or the link that caused the breach (*e.g.*, whether her credentials were stolen by keylogger, phishing, or brute-froce *etc*).

For the outcome to be in doubt Alice must have unlimited budget: she does whatever it takes to protect her resources against Charles. However, if Alice has a fixed budget, her situation appears hopeless. How can she defend against an enormous and ever-growing, ever-adapting set of attacks? In this threat model, failure to do everything means that there is no point in doing anything. This leads to the situation of all-or-nothing investment: unless both A and B are done the investment in either is wasted. This is particularly problematic for end-users who clearly have limited resources, and yet are subject to an extraordinary array of internet attacks. Alice's correct strategy is "do everything." If that is not possible, she may as well purse the alternative strategy "do nothing and hope for the best." In this threat model there is little point in any of the strategies in between.

In the model where Charles exploits any vulnerability, the worst-case is to be expected if any defence is neglected. This leads to the following puzzling fact: the idea that worst-case outcomes become actual is not supported by evidence [6]. In spite of the huge and growing list of attacks, close to two billion people use the internet regularly for email, banking, social networking and a host of other activities. Bad things certainly happen, but apparently not often enough to outweigh the benefit that people derive from using the internet and working online. We need not look to user diligence for an explanation of this contradiction. Not only do users take no precautions against elaborate attacks, they appear to neglect even basic ones. For example, a growing body of measurement studies make clear that users choose weak passwords and re-use them liberally [21], choose easily guessed backup authentication questions [43], are oblivious to security cues [41], ignore certificate error warnings [26] and cannot tell legitimate web-sites from phishing imitations [19]. How can this be? How can it be that huge numbers of internet users ignore most security advice, and yet do not suffer the worst-case outcomes predicted by this threat model? Alice, clearly presents no shortage of vulnerabilities, and yet Charles somehow fails to exploit them.

The contribution of this paper is to suggest that an explanation lies in basic economics. We do so by proposing a new threat model. We split Charles' attack efforts into those of two attackers: Carl and Klara. Carl mounts only scalable attacks (where the costs grow slower than linearly in the number of users attacked) all other attacks are carried out by Klara. Examples of scalable attacks are those that are automated, *e.g.*, phishing and spam. Examples of non-scalable attacks are those that involve per-user effort, *e.g.*, spear phishing.

We show that this very simple cost restriction has profound implications for the nature and range of Carl's attacks, and how Carl and Klara divide the attack opportunities between them. In summary, Carl's attacks are large, broadcast internet-scale attacks. For the same cost, Carl attacks orders of magnitude more users than Klara. Carl's economies of scale force Klara to be extremely selective. She must target resources where value is very concentrated, so that she can extract maximum value from very few users. This means that while everyone is attacked by Carl, very few are targeted by Klara, even when they are vulnerable. Just as misaligned incentives for defenders can produce unexpected results [37], the lack of clear return on effort for Klara can cause real vulnerabilities to go unexploited. This also has implications for Alice's resource allocation decisions. Alice must protect first against all scalable attacks. Since these are large, visible, predictable and non-adaptive she has some hope of accomplishing this with a limited budget.

# 2. SCALABLE ATTACKS

## 2.1 Threat Model: Carl and Klara

In a refinement of the Threat Model where Alice faces a single attacker, we propose one where she faces two: Carl and Klara. Rather than restrict his technical capabilities Carl is constrained in that he can mount only scalable attacks (defined below), everything else is the province of Klara. There is no loss of generality here: no attacks are ruled out of scope, we merely label those that are scalable as belonging to Carl and all others as belonging to Klara. Carl can collude with Klara so long as it doesn't violate the constraint of keeping costs scaleable. One obvious form of collusion is a flow of information between them: *e.g.*, Carl passes prospects for non-scalable attacks to Klara if he can identify them using a scalable attack (as in Section 4.4). We will see that their different cost structures have major implications for the types of attacks that Carl and Klara mount and how they divide attack opportunities between them.

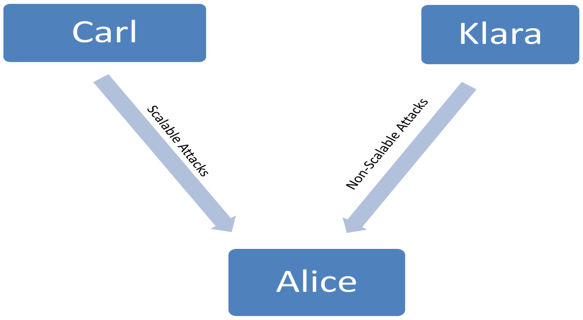

**Figure 1: Threat model b: the attacks on Alice are scalable (from Carl) and non-scalable (from Klara).**

We address the case where Alice is an end-user. This simplifies the analysis, since there is a single decision maker and the assets she protects are her own. Alice has several assets that are of interest to attackers. She has bank, email, social networking accounts; she has a computer which provide bandwidth, computation and hosting.

## 2.2 Scalable Attacks: Carl

The term scalable is used without a standard definition in networking, economics and business planning. For the purposes of this paper an attack is scalable if the cost, $C_s(N)$, grows slower than linearly in the number of attacked users, $N$ :

$$C_s(2N) < 2C_s(N). \qquad (1)$$

This means that the attack achieves economy of scale (*i.e.*, costs decrease with increasing "customers") once the startup costs have been recouped.

This cost model constrains Carl. Attacks that involve individual per-user effort or cost are ruled out. Equally, requiring physical proximity to, or knowledge of, the user is not possible with a scalable cost structure. It might appear that Carl can mount only a small range of very constrained attacks. However, these are some of the most widely deployed web and internet exploits. Examples of scalable attacks are spam, phishing, viruses, drive-by downloads, self-replicating botnet code *etc.* Costs for each of these clearly satisfy (1).

### 2.3 Non-Scalable Attacks: Klara

Non-scalable attacks are everything else. So the cost grows linearly, or worse:

$$C_n(2N) \approx 2C_n(N). \qquad (2)$$

Each additional attacked user costs as much as the last.

Cost structures of this kind clearly occur if dedicated effort has to be expended on each attacked user, if physical proximity to the user is required, or if particular knowledge of the attacked user is needed. An example of a non-scalable attacks is spear phishing, where a personalized phishing email is sent to each recipient [32]. Attacks that seek to spy using reflected light, electromagnetic emissions, or sound [29, 31, 34] clearly require proximity and are non-scalable. Session-hijacking attacks that involve real-time Man-In-The-Middle activity on a live session clearly involve per-user effort. Certain attacks on the back-up authentication questions [43] clearly involve both dedicated knowledge of the user and per-user effort. Most social engineering attacks involve dedicated time-consuming effort and are non-scalable. Almost all of the exploits described by Mitnick [27] involve dedicated effort and are non-scalable.

### 2.4 Mathematical Model

It is clear from experience that many internet attacks are run at such large scale, and are so wasteful, that costs don't merely grow slower than linearly, they barely grow at all: $C_s(2N) \approx C_s(N)$. It would be an understatement to say, for example, that spammers' costs do not have a dependence on $N$ (forcing such a dependence has been suggested as a solution to automated attacks such as spam [12]). In fact, many scaleable attacks resemble information goods or software businesses in the sense that the first-copy costs dominate, and subsequent copies are almost free [15]. For example, consider the spam campaign documented by Kanich *et al.*[28], were a spammer made \$2800 from 350 million emails sent. If we assume that the spammer (Carl) at least broke even, then $C_s(350e6) = \$2800$ for this attack. By contrast, if Klara devoted even \$0.10 of effort per-user, her cost to reach the same population would be \$35 million. Alternatively, for the same cost, Klara could instead attack 28k users (*i.e.*, four orders of magnitude fewer). If she spends \$1 per user it grow to five. If Klara spends an hour of US minimum wage effort per user she gets to attack $2800/7.25 = 386$ users for the same cost that Carl attacks 350 million, a six orders of magnitude difference.

We've stated the costs, what of the rewards? We'll assume that, for both Car and Klara, the rewards are $R(N) = NY\overline{V}$, where $N$ is the number of attacked users, $Y$ is the yield, and $\overline{V}$ is the average *extracted* value per *successfully* attacked user. Thus $N$ are attacked, $NY$ are successfully attacked and $NY\overline{V}$ is the value extracted.

| | Scalable | Non-scalable |
|---|---|---|
| Cost | $C_s(2N) < 2C_s(N)$ | $C_n(2N) = 2C_n(N)$ |
| Reward | $R_s(2N) = 2R_s(N)$ | $R_n(2N) = 2R_n(N)$ |
| Profit | $P_s(2N) > 2P_s(N)$ | $P_n(2N) \leq 2P_n(N)$ |

**Table 1: Dependence of Cost, Reward and Profit on number of attacked users $N$. We merely constrain that the costs of scalable attacks grow slower than linearly; *i.e.*, the second derivative is negative. We call all other attacks non-scalable.**

## 3. ANALYSIS AND IMPLICATIONS

### 3.1 Scalable Attacks Run at the Largest Possible Scale

A consequence of the slower than linear cost growth is that there is a powerful incentive is to run scalable attacks at the largest possible scale. For Carl, profitability increases monotonically with the number of attacked users:

$$\begin{aligned} P_s(2N) &= R_s(2N) - C_s(2N) \\ &> 2R_s(N) - 2C_s(N) \\ &> 2P_s(N). \end{aligned}$$

It's more profitable to attack $2N$ users than $N$, and better to attack $4N$ than $2N$ and so on. Scalable attacks should be limited only when no more possible victims can be identified. Hence Carl has the incentive to attack everyone as often as possible. This correlates well with experience that nobody is immune from spam, phishing and other large scale scriptable attacks.

By contrast for Klara the profit picture is very different:

$$\begin{aligned} P_n(2N) &= R_n(2N) - C_n(2N) \\ &= 2R_n(N) - 2C_n(N) \\ &= 2P_n(N). \end{aligned}$$

She certainly sees no improvement in profitability with scale. At best, she makes as much from attacking her million-and-first user as she did from her first. Since Carl attacks everyone Klara always competes with him. Unless she can find an asset class that is not subject to scaleable attack everyone she targets is also attacked by Carl. However, Carl has far greater reach. In Section 2.4 we saw that a \$0.10 per-user cost would cause Klara to reach $N_n/N_s = 10^4$ fewer potential victims than her spamming rival for the same cost.

### 3.2 Scalable Attacks are Automated

Consider a conventional spam campaign, with cost $C$. The expected value of a victim is $\overline{V}_s$, and $N_s$, $Y_s$ are the number of contacted recipients and yield respectively. If the campaign is profitable then:

$$C_s(N) < N_s Y_s \overline{V}_s.$$

Now consider personalizing the campaign: the sender makes the effort to include some information identifying the recipient in the mail. For example, instead of beginning "Dear Respected Sir" it begins "Dear Mr. Brown" or "Dear Julie Smith" or has other customization that cannot be performed automatically and involves an additional cost of $\beta$ per user. For example, an Internet search per user to add some personalizing information would add such a cost. The yield presumably increases because of the personalization, $Y_s^{'} > Y_s$, so the return also increases to $N_s Y_s^{'} \overline{V}_s$. This is an improvement over the un-targeted attack, so long as the improvement in reward per user is greater than the additional cost:

$$\beta < (Y_s^{'} - Y_s) \cdot \overline{V}_s.$$

Let's suppose that targeting increased the yield by 4.5, *i.e.*, $Y_s^{'} = 4.5 \times Y_s$ (this is the factor improvement that Jagatic *et al.* [46] found in a customized phishing attack over a generic one). Using the numbers from the spam campaign documented by Kanich *et al.* [28]: $N_s = 350e6, Y_s = 28/N, \overline{V}_s = \$100$. This gives that, to be an improvement, the targeted campaign must have $\beta < \$0.00002$. Using the US minimum wage of $7.25 an hour this translates to 0.01 seconds effort per-user.

This arithmetic is orders of magnitude away from making economic sense. While personalization would certainly improve yield, the economies of scale of Carl's scalable attack model overwhelm any advantage this might bring. Thus, scalable attacks must be entirely automated with no per-user intervention whatever. This matches the experience that we have from spam, phishing *etc.* The majority of spam and phishing emails contains no personalization whatever; Nigerian 419 scams seldom contain identifying information.

## 3.3 Scalable Attacks don't Adapt to Individual User Actions

Carl cannot personalize his attacks for Alice. Equally, he cannot adapt to her actions and counter-measures: anything that involves effort customized to an individual violates his scalable cost structure. This forms a key point of contrast between the attacks of Carl and Klara. For example, Carl and Klara both want Alice's bank password, and both have a range of attacks to try to get it. Carl, for example, will try phishing, will send spam with keylogging Trojans, he will probe Alice's firewall for un-patched vulnerabilities, and perhaps check if her router still has the default password. These are all scalable attacks that fall within Carl's range. Klara may try guessing the backup authentication questions, she may try more elaborate social engineering based on what she can learn of Alice [27, 46], she may physically instal a keylogger on Alice's machine, or even snoop using the audible [29], visible [34] or electromagnetic spectrum [31].

A key difference is the non-adaptive nature of Carl's attacks. If Alice evades the phishing and emailed Trojan, Carl doesn't step up other attacks. There is no increase in his attacks on the router if he is defeated at the firewall. By contrast Klara adapts: if attempts on the backup authentication questions fail, she may try harder at more advanced social engineering. If these again fail she may endeavor to get a keylogger installed. If Klara is truly determined and possesses unlimited resources, Alice faces a daunting challenge. If Klara is a family member or coworker, physical access to Alice's computer may not be a problem, and installing a keylogger may be perfectly feasible. If Klara works for a state agency breaking and entering may be possible. We assume that both Carl and Klara are rational. Since Carl looks at all users indiscriminately his spending is constrained by $C_s(N) < N_s Y_s \overline{V}_s$. He won't spend more on Alice (or any other user) than the average expected return. Klara, on the other hand, may persist, especially if she believes that she can extract more than $\overline{V}_s$ from Alice. We explore this in Section 4.2.

In fact, not only does Carl not adapt when he fails, he doesn't adapt when he succeeds. If he phishes Alice's PayPal account he continues to try to phish her; even if she has no money left, even if she closes the account. It is also for this reason that Carl attempts to instal botnet code, even if several other Trojans are already running. This is an consequence of the automated nature of the attacks. Carl doesn't know his potential victims; he doesn't even know how many of those he attacks are real people. This is one reason for the success of honeypot techniques against scaleable attacks [36].

## 3.4 Scalable Attacks Produce Commodity Goods

The cost structure of scalable attacks is both their strength and their weakness. Since these attacks are automated they can be replicated without much skill. A single clever exploit, once scripted, can be used by many. These "script kiddie" attackers need not necessarily understand in detail how the attack works or have very much skill.

Once an attack is automated it is difficult to preserve profit. The script or kit can be passed to many: since it is automated anyone can do it. Any insight or intellectual property involved is diffused to many. If anyone can do it, more and more people are attracted to the opportunity so long as it is profitable.

As in other industries the effect of automation is commoditization of the product. For Carl the product may be spam delivered into inboxes, bank passwords stolen, or computers enlisted to serve in botnets. Automation greatly increases supply and drives the price down.

Zombie PC's, for example, become less valuable if large botnets are common. If bank credentials are monetized through a mule, and mules are in short supply then increasing supply of credentials drives the value down. This appears to be the case with phishing where, despite very visible effort there is slender evidence of return [24]. A similar effect is visible in spam, where enormous campaigns appear to generate minuscule returns [28]. A consequence of the Tragedy of the Commons is that returns drop with increasing effort. That is to say, $\overline{V}_s$ decreases. For example, the value of getting into a user's email inbox has probably decreased with time. Since so many attackers attempt to phish, there is evidence of greater supply of credentials than can be successfully be harvested [25]. There is evidence that the value of zombie PC's has been falling with time [4].

This is in line with history of industrialization. Mass production greatly increases supply and turns the product into a commodity, which trades primarily on price [35]. There is little opportunity for producers to add value, and the advantage goes to the lowest cost producers. Again the situation resembles an information good where there is no barrier to entry; according to Shapiro and Varian "competition among sellers of commodity information pushes prices to zero" [15]. The existence of a scaleable attack on a particular asset turns it into an economies of scale business.

## 4. SCALABLE VERSUS NON-SCALABLE

We now explore the question of how Klara and Carl divide the attack opportunities between them. For the same cost, $C$, Carl and Klara attack different numbers of users. Thus, Klara reaches only a small subset of Carl's population. That is, when $C_s(N_s) = C_n(N_n) = C$, Carl's audience, $N_s$, exceeds Klara's, $N_n$, by orders of magnitude. Now, if she is is to match Carl's rewards Klara needs:

$$N_n Y_n \overline{V}_n \geq N_s Y_s \overline{V}_s.$$

This implies the following inequality between the ratios of their yields, reaches and extracted values:

$$\frac{Y_n}{Y_s} \geq \frac{N_s}{N_n}\frac{\overline{V}_s}{\overline{V}_n}. \tag{3}$$

So she requires:

$$\log_{10}\frac{Y_n}{Y_s} \geq \log_{10}\frac{N_s}{N_n} - \log_{10}\frac{\overline{V}_n}{\overline{V}_s}. \tag{4}$$

This constraint is shown pictorially in Figure 2. This shows the region where Klara beats Carl for several different values of $N_s/N_n$. When Carl outreaches her by $N_s/N_n$ then Klara beats his return above and to the right of the Yield-Value frontier shown in Figure 2. Since $N_s \gg N_n$, she has two directions she can explore to offset his advantage in numbers: she can try to achieve yield that exceeds his, or target subsets of the population where she can extract far greater value than his average $\overline{V}_s$, or some combination of the two.

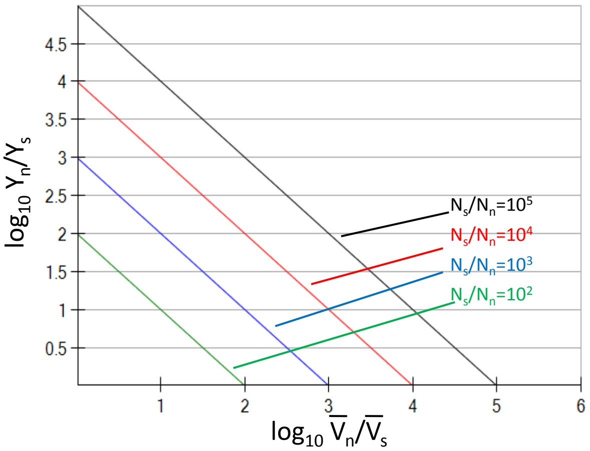

**Figure 2: The profit frontier at which Klara's targeted attacks beat Carl's scalable attacks. When she is outreached $N_s/N_n$ she needs the ratio of her yield, $Y_n/Y_s$, and average extracted values, $\overline{V}_n/\overline{V}_s$, to exceed Carl's by at least the amount shown. For example, if $N_s/N_n = 10^4$, if she can achieve $\overline{V}_n/\overline{V}_s = 10^3$ then she beats Carl so long as $Y_n/Y_s > 10$. Above and to the right of the profit frontier she does better than Carl, below it she does worse.**

### 4.1 Competing on Yield Alone Makes No Sense

If she competes on yield alone Klara effectively competes with Carl on price. That is, if she extracts the same value per victim (*i.e.*, $\overline{V}_n = \overline{V}_s$) then she must have a lower cost per successfully attacked user. There are a few circumstances where $\overline{V}_n = \overline{V}_s$. First, this happens if the distribution of extractable value is uniform; *i.e.*, all user's yield equal value when successfully attacked. Second, the distribution of extractable value is unobservable; *i.e.*, some users have higher value than others, but there is no way of determining which. In each of these cases Klara cannot hope for higher extractable value and must compete on yield alone.

When $\overline{V}_n = \overline{V}_s$ Klara's constraint (3) simplifies to:

$$\frac{Y_n}{Y_s} \geq \frac{N_s}{N_n}. \tag{5}$$

Thus, to match Carl's return, Klara's yield must be a factor $N_s/N_n$ greater than Carl's. She must exceed his yield, by as large a factor as he exceeds her reach. This may be possible, depending on Klara's skill. In part because of the competition discussed in Section 3.4, yield on scalable attacks seems to be low in general. For example, Kanich *et al.* [28], reported a yield of $3.8\times10^{-6}$

for infections from greeting card spam, and $8 \times 10^{-8}$ for pharma spam. Yields this low for Carl do hold open the possibility that, while hard, Klara may be able to make up with yield what she loses with reach (of course if $Y_s > N_n/N_s$ then even perfect yield, $Y_n = 1$, wouldn't allow Klara to match Carl). On the other hand, Jagatic *et al.* [46] report 16% success with generic phishing emails and a 72% success rate with targeted spear phishing emails. The 4.5× improvement with personalization is of little help to Klara if she has a $10^5$ deficit in reach to make up.

Recall, from Section 3.4, that (for those assets that Carl attacks) $\overline{V}_s$ decreases with time. Thus, when $\overline{V}_n = \overline{V}_s$, both Carl and Klara see their returns decrease. This is an effect of the competition that exists when an attack is automated as Carl's are. Since he does not have a fixed cost per user, Carl can respond by increasing $N_s$ or launching increasing numbers of attacks. Since Klara does have a per-user cost she cannot respond in this way. Thus her cost per user is fixed, while the extractable value per user is driven down by the commoditization that Carl causes.

To summarize, competing on yield alone is the same thing as competing on price for Klara. This is unpromising for a number of reasons. First, there is little evidence that personalization can make up for the orders of magnitude deficit she suffers in reach. Second, Carl turns the assets he attacks into commodities, and drives their prices down over time. Even if she competes successfully her situation gets worse with time. Her values decline, while her costs are constant. Finally, history suggests that competing with a mass-production competitor requires mass-production. Just as corner stores get driven out by chains, and chains get driven out by larger chains, an economies of scale business is unkind to participants who can't scale. History also suggests that the way out is to seek higher value niche opportunities. This has been the pattern in scores of industries: those who have a linear cost model must seek the highest value part of the market and differentiate their product from the commodity version. Thus Klara must seek users with higher extractable value. Ideally $V(k)/\overline{V}_s$ is large, but at the very least, she must seek users with higher than average value: $V(k) > \overline{V}_s$. Only in this way can she avoid competing in a commodity space.

## 4.2 Seeking Higher Value Targets

Let's sort all users by extractable value $V(k)$ giving a decreasing function as shown in Figure 3. Since she can reach only a fraction of the population, it is clear that Klara's return is best when she targets the users with highest extractable wealth. At the very least, we have seen that she needs users with higher than average value.

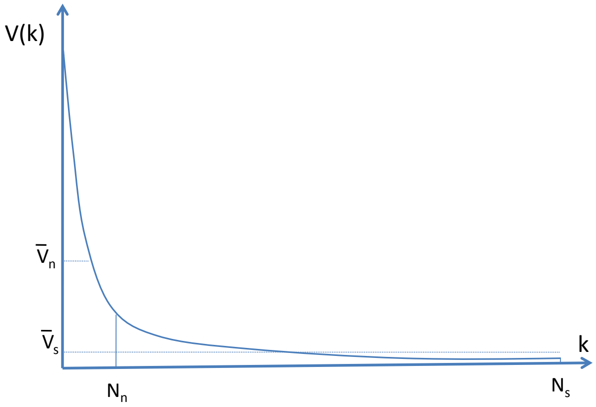

**Figure 3: Instead of attacking the whole population Klara concentrates on the most valuable segment. She targets those with highest value, leaving the remainder to Carl. The average value of her targeted users is $\overline{V}_n$. For concentrated distributions the average is higher than the median; so far fewer than half of users have greater than average value.**

### 4.2.1 *Klara Needs Longtail Distributions*

To avoid competing with Carl on cost, Klara needs to target users with higher than average value: $V(k) > \overline{V}_s$. This is easiest when extractable value is concentrated among as few users as possible. To profit from this concentration she also requires that it be observable which users have highest value. Thus she needs:

- Concentration of extractable value
- Visibility of extractable value.

If she can attack only $N_n$ users, she wants $\overline{V}_n = 1/N_n \sum_{k=0}^{N_n-1} V(k)$ to be maximum. That is, she needs as big a fraction as possible of the total value to be concentrated among as few users as possible. This ensures that $\overline{V}_n/\overline{V}_s$ will be as big as possible (which assists with the goal of being on the right side of the profit frontier in Figure 2). A distribution of extractable value $V(k)$ that is uniform is the worst case for Klara: she can do no better than attack indiscriminately. A distribution that is unobservable is no better: there may be high value users, but she can't figure out which ones they are. Best for her are skewed distributions with long tails such as exponential, and power-law distributions such as Zipf, Pareto *etc.* These distributions (an example is shown in Figure 3) generally have considerable concentration: some users have value much higher than others, which is precisely what Klara needs. The greater the concentration the higher $\overline{V}_n/\overline{V}_s$, and the better the chance of lying beyond the profit frontier of Figure 2. In this she is fortunate: power-law distributions are common in many naturally occurring phenomena. The distribution of wealth [20], income, fame [14], size of human settlements, word frequencies, web-site populatritics and many other phenomena are well modeled by very concentrated power-law distributions.

### 4.2.2 In Longtail Distributions Most Users Have Below Average Value

One might imagine that to find users with $V(k) \geq \overline{V}_s$ Klara need merely target the top half of the population. However, monotonic distributions with positive skew, such as exponential, Zipf, Pareto *etc*, have the property that the mean is greater than the median. For example, the average wealth (or income or house price) is generally higher than the median, since a few very large samples pull the average upward. This means that far fewer than half the people have the average wealth. Thus, if value is concentrated (as Klara requires) fewer than 50% of users have $V(k) \geq \overline{V}_s$. The more concentrated the distribution the greater the difference between the mean and the median, and the fewer people who exceed the average.

For many naturally occurring power-law distributions the concentration is dramatic. Using the US wealth distribution example [20]: only 1.8% of people exceed the average wealth. In this instance more than 98% of people have below average value $V(k) < \overline{V}_s$. Attacking these people would hurt rather than help Klara's returns. In a study of literary fame Martindale [14] discovered that half of the scholarly attention was devoted to a mere 2% of the English poets considered. In a study of scientific productivity Price [18] suggested that in a discipline with $N$ scientists half of the papers are produced by $\sqrt{N}$ of the scientists (and thus concentration increases with the size of the discipline). In each of these cases a very few percent at the top account for half of the value. Now, to satisfy (4), and beat Carl, Klara needs $\overline{V}_n \gg \overline{V}_s$. It certainly doesn't help to include in her target population users for whom $V(k) < \overline{V}_s$. So the number of users who exceed average value is a loose upper bound on the number that Klara wishes to attack.

Table 2 shows the number of users with above average value for a number of distributions. They do make clear that in the concentrated distributions only very small fractions of the population have greater than average value. Thus, the vast majority of the population is not of interest to Klara. For example, if $V(k)$ follows the same distribution as fame then the least valuable 98% of the population is worthless to her: attacking them merely reduces her return.

One might imagine that the fact that Klara reaches orders of magnitude fewer users than Carl establishes than most users experience no targeted attacks. Indeed, if she had the field to herself, Klara should attack only the top $N_s/N_s$ fraction of the population. When Klara is outreached by $10^4$ this suggests that Klara focusses on the top 0.01% of the user population. However, this is only the case when a single scalable and a single non-scalable attacker make equal investments. When there are many different targeted attackers the attacks may spread to more and more of the population. However, no matter how many targeted attackers there are, no matter what the investment in non-scalable attacks, in no case does attacking those with below average value make sense. Thus, when extractable value is concentrated, most users are not profitable targets for non-scalable attacks.

| Distribution | $V(k) > \overline{V}_s$ |
|---|---|
| Wealth US [20] | 1.8% |
| Fame [14] | 2.0% |

**Table 2: Percent of users with above average value for various distributions. This is a loose upper bound on the fraction of users that Klara targets. Attacking those with below average value reduces her return.**

## 4.3 What Assets Does Klara Target?

The best strategy for Klara involves going after resources where the distribution of value is both concentrated and observable. This is far from a simple. Extractable value is not necessarily related to actual wealth: a user who has net worth of $1 million doesn't necessarily have extractable value 100× greater than one with $10k. It's unclear, for example, that Klara would extract much if she set about attacking Bill Gates or Warren Buffet.

Which types of assets offer the visible value concentration that Klara requires? First, an example of an asset that is probably bad for Klara is attacking PC's to enlist in a botnet. If we set aside the question of any credentials on a machine, computers probably have too close to uniform value when enlisted in a botnet. That is, while there will be variation in value (*e.g.*, some computers are faster, with higher bandwidth and good IP reputation *etc*) it is unlikely that a small segment with higher than average value can be identified. Taking the credentials into account alters the picture: a PC that is used frequently for online banking might have value much greater than one that is not, as the credentials add value. But it is not observable to an attacker which PC's contain credentials until after an attack has succeeded. So the observable part of the value appears close to uniform, while the concentrated part is unobservable.

Second, email and social networking accounts may be very good targets for Klara. The average email account has little extractable value, *i.e.*, $\overline{V}_s \approx 0$. For example, a hotmail account might be used to send spam and harvest contacts, but only in rare cases will it yield large value. However when it belongs to someone famous an email account is worth a great deal, *i.e.*, $V(k) \to \infty$. Sarah Palin's Yahoo! email account clearly had value many times greater than the average. Equally, the email records of the University of East Anglia climate scientists had great value to climate change skeptics who wished to discredit their findings. A similar factor was probably at work when security researcher Dan Kaminsky's email and server data were compromised shortly before the Blackhat conference in 2009. Famous people are ideal targets, since fame is both very concentrated and very observable. Further, value extraction is not a problem. Leaking Palin's emails was the sought-after prize, so no error-prone monetization strategy was required. These are precisely the circumstances that favor Klara: very visible concentration of value. A further advantage, is that when $\overline{V}_s \approx 0$ she faces very little competition from Carl.

Third, what of bank credentials? Since phishing is one of Carl's favorite exploits there may be little point for Klara in seeking bank credentials indiscriminately. Carl is already achieving a high enough yield with a scalable attack. In fact, there have been numerous accounts that Carl harvests more accounts than can be successfully drained: bank credentials are offered for fractions of a penny on the dollar [3, 40, 25]. There is little point using a non-scalable attack and dedicated effort per-user to steal credentials, when there's already a surfeit of them with asking price \$2 on the underground economy [3]. However, concentration of wealth creates favorable conditions for Klara if she can identify cases where extractable value is highest. Krebs [1] has reported a number of high value attacks on small businesses. Small businesses are an excellent avenue for Klara. Checking accounts with far larger amounts than consumers would possess but with relatively weak protections may produce the happy combination of $Y_n/Y_s$ and $\overline{V}_n/\overline{V}_s$ that lies above the profit frontier of Figure 2.

There are cases where value may be concentrated, but is unobservable. This is often the case if the higher value of an asset is related to bad security practices. For example, some users have the same password at hotmail as at their bank. The passwords of these hotmail users are far more valuable than the average. However, the higher value appears to be unobservable: there's no simple way to identify which users follow this practice. Equally, computers that are used for online banking, or that have lots of stored credentials have higher value, but this is not observable.

Finally, a special case of particular importance is when Klara values the resources of some users much higher than any other attacker would. This is the case when she knows her targets. Control of the email account of a randomly chosen internet user, Alice, is of almost no value to a randomly chosen attacker. However, to those who know and are close to Alice it may have great value. A suspicious husband, jealous ex-boyfriend, curious sister or ill-intentioned roommate might value access to her email a great deal. For an attacker who cares greatly about Alice's resources and very little about anyone else's: $V(\text{Alice})/\overline{V}_s \to \infty$. In this case the concentration of value is almost infinite and visibility is perfect; so, it is worthwhile attacking no matter what the effort. Closeness increases the value, proximity and access improve the yield.

## 4.4 Scalable Recruitment for non-scalable Attacks

Some attacks have an an scalable recruitment phase and then non-scalable follow-up. The Nigerian 419 scam recruits prospects using a spam campaign; those who respond receive individual attention. Mule recruitment also follows this pattern: an advertising campaign for a "work at home" opportunity, is followed by individual attention for those who respond. The attacker casts as wide a net as possible in the scalable phase before following up with the expensive linear-cost phase. One way of viewing these attacks is that those who are susceptible to Nigerian scams or mule schemes have higher, but unobservable, extractable value. A scalable attack from Carl serves merely to reveal those who are worth targeting. Klara follows up with a non-scalable attack on the newly-revealed high value targets.

## 4.5 On the Internet Nobody Knows You're Not a Dog

The resolution of the paradox we posed in the introduction appears very simple. To avoid competing with Carl on price, Klara needs users with above average extractable value. She also needs value to be concentrated among as few users as possible. For the distributions that give the kind of concentration that she needs very few users have above average value. For the power-law, longtail and 80-20 type distributions that model many natural phenomena fewer than 1% of users will have above average value. Thus 99% of users are not targeted by Klara; attacking them would hurt rather than help her overall return.

Thus, the reason users escape disaster, even when they flout security advice, is that the vast majority of the online population is attacked constantly by Carl, but never targeted by Klara. Their bank accounts may be hack-able with a few hours of effort. But Carl won't spend the time, since it violates his cost model. Klara won't spend the time, unless she knows that Alice's account has extractable value higher than the average.

For example, consider the case where the backup authentication questions to Alice's bank account can be determined with one hour of effort from publicly available information (*e.g.*, her facebook page *etc*). Further suppose that the account has $200 extractable value. We might regard this an opportunity to make $200 for an hour's work. But to do so ignores the survivor paradox. Klara must *know* that the effort yields this return for this effort. If Klara succeeds only one time in a hundred with this approach her average return drops to $2 an hour, if one time in a thousand it drops to $0.2. Thus, the fact that the profitability of the target (*i.e.*, $200 for an hour's work), is obscure to Klara, saves Alice. Alice doesn't get security from this obscurity, but the fact that the cost of removing the obscurity is greater than the expected return, helps her escape harm. In this case Alice's avoidance of harm is determined not primarily by her security measures, but by the relative worthlessness of the average facebook page, and the fact that her higher value is unobservable. On the Internet nobody knows you're not a dog.

### 4.6 Allocating Resources: Most Users can Ignore Most Attacks

This analysis also suggests a resource allocation strategy for Alice. In common with Bart *et al.*[7] and Boehme and Moore [39] our conclusions suggest that under investment in security can be rational. Our analysis suggests that most people will not experience any of the non-scalable attacks that involve per-user effort. If the distributions of value are as concentrated as those of wealth and fame then fewer than 2% of users will ever see such attacks. If the distributions are not concentrated then it is almost impossible for Klara to make up for the orders of magnitude greater reach that Carl enjoys.

Thus, all users should protect against scalable attacks first. Compromise is almost certain if Alice fails to address the scalable attacks that reach everyone. After this, Alice's strategy depends on which, if any, of her assets are valuable enough and visible enough to place her in the top few percent of available targets. Only a small fraction of users can expect to see high cost attacks. At one extreme attacks that involve light [34], sound [29] or electromagnetic [31] emanations probably affect a minuscule number of users, and can safely be ignored by most. Less extreme are attacks that involve passwords that are written down or re-used, or answers to backup authentication questions that can be determined from public information. These attacks are largely non-scalable. While each opens a security vulnerability of some form, the vulnerability is unlikely to be exploited unless Alice offers high extractable value that is visible to an attacker. Even using her dog's name as password, while hardly advisable, increases risk only from an attacker who is already targeting Alice. It makes sense that Alice not spend more on defense than an asset is worth. However, our analysis suggests that in addition she can neglect defending against attacks where the cost is greater than the expected value (to the attacker) of the asset.

While new attacks are discovered regularly, new scalable attack vectors are comparatively rare. Phishing persists even though it is a relatively old attack and many users must be familiar with it. In spite of its age spam is still one of the most common methods for delivery of malware. The very special cost structure of scalable attacks guarantee that most new attacks are non-scalable.

## 5. RELATED WORK

The question of tradeoffs in security is not a new one. Numerous authors have pointed out that, even though security is often looked at as binary, it cannot escape the budgeting, tradeoffs and compromises that are inevitable in the real world. The scalable nature of many web attacks has been noted by many authors, and indeed this has often been invoked as a possible source of weakness. Dwork and Naor [12] first suggest addressing spam (and other resource-abuse attacks) by forcing a linear dependence between use of the resource and cost. The importance of the difference between linear and sub-linear costs for an attacker has been recognized by others who have attempted to prove or disprove the feasibility of this approach [9, 17].

There have been numerous studies documenting the enormous range of internet attacks. Sariou *et al.*[42] perform an interesting measurement study of internet attacks. Kanich *et al.*[28] document the result of observing a spamming botnet for a number of weeks. Their findings provide interesting insight into the scale and yield of scalable attacks. Prior to their work, we have had surprisingly little data on the cost and scale of spam campaigns. Stone *et al.*[11] also managed to take over a botnet for a period of weeks. Phishing has been the focus of a few measurement studies. Florêncio and Herley [21], using password entry data from toolbar users, estimate the annual phishing victimization rate at 0.4%. Using the entirely independent method of counting credentials at compromised phishing servers Moore and Clayton [45] estimate it at 0.34%.

Since 2002 the Symposium on Usable Privacy and Security has investigated tradeoffs between security and usability, especially where end-users are concerned. Several authors in the Usable Security literature have drawn attention to the difficulty of the resource allocation situation that end-users find themselves in. Adams and Sasse [8] demonstrate that users have difficulty with simple security policies and have poor understanding of the risks. Herley [23] argues that users are correct to ignore most of the security advice they receive: faced with costs which exceed their budgets and benefits that are not evidence-based they have little alternative but to capitulate and hope for the best. Beautement *et al.*[5] introduce the idea of a compliance budget for users, suggesting that once the burden of security policies reaches a certain point further complexity becomes intolerable. The resource allocation we suggest in Section 4.6 can be seen as a way of prioritizing when the budget is fixed. Anderson [37] shows that incentives greatly influence security outcomes. Our work suggests that the lack of clear return for non-scalable attacker can explain the failure to exploit many weak links. Since 2000 the Workshop on the Economics of Information Security (WEIS) has focussed on incentives and economic tradeoffs in security.

Jackson [13] suggests that the web attacker needs three things: bad code, a means to get it running, and an introduction to the user. It is often the last of these three, the introduction to the user, that forms the scalable part of Carl's effort. He spams as many email addresses as he can find or guess. He port-knocks entire ranges of IP addresses. He tries to lure as many users as possible to his vulnerability-exploiting web-site.

Barth *et al.*[7] examine the question of reactive security, and show that it can be effective in settings where the defender does not myopically over-react to the most recent attacks. While the theoretical framework is rather different, our finding for Alice does echo this result. We suggest that a reactive approach to Carl's attacks and proactive to those of Klara only when necessary is the best use of defender effort.

Odlyzko [6] addresses the question of achieving security with insecure systems, and also confront the paradox that "there simply have not been any big cybersecurity disasters, in spite of all the dire warnings." His observation that attacks thrive in cyberspace because they are "less expensive, much more widespread, and faster" is similar to our segmentation of scalable attacks. Schneier [2] argues that "one of the important things to consider in threat modeling is whether the attacker is looking for any victim, or is specifically targeting you." This is one of the conclusions we draw: everyone is attacked, but very few are targeted. Those who are valuable enough to be targeted face a very different climate from the majority who are not.

Varian suggests that many systems are structured so that overall security depends on the weakest link [22]. Gordon and Loeb [30] describe a deferred investment approach to security. They suggest that, owing to the defender's uncertainty over which attacks are most cost effective, it makes sense to "wait and see" before committing to investment decisions. Boehme and Moore [39] examine an adaptive model of security investment, where a defender invests most in the attack with the least expected cost. Interestingly, in an iterative framework, where there are multiple rounds, they find that security under-investment can be rational until threats are realized. Unlike much of the weakest-link work, our analysis focusses on the attacker's difficulty in selecting profitable targets rather than the defender's difficulty in making investments. However, investment strategies that echo the findings of Gordon and Loeb [30] and Boehme and Moore [39] spring from our findings.

## 6. CONCLUSION

The important distinction between attacks that are scalable and those that are not has long been recognized. Dwork and Naor [12] suggest that forcing a linear cost dependence makes certain attacks unattractive. The scalable attacker lives in a "costs nothing to try" world. He attacks everyone indiscriminately, whether they have high value or low, and achieves the average value per user. By contrast, every attacked user is a real expense to the non-scalable attacker. Since she reaches orders of magnitude fewer users she must choose her targets with care as low value users hurt her return.

If her returns are to match the scalable approach she needs value to be both concentrated and observable. Favorable circumstances for her are longtail and power law distributions of value, where a small number of users have very high value. This value must also be visible. However, in these distributions only a tiny fraction of users have above average value. Thus, the circumstances that favor Klara are exactly those that ensure that she attacks very few users. Extreme concentration of value, such as is the case with wealth and fame, will ensure that she attacks only a percent or two of users.

This helps explain the fact that worst-case outcomes fail to materialize: the cost of non-scalable attacks is such that very few users are targeted. It further suggests a security investment strategy for Internet users: all scaleable attacks should be addressed first. Consider the case where Alice's email account can be harvested for value $200 by a non-scalable attacker. Alice's avoidance of harm depends not so much on her security investments, but on the relative worthlessness of other email accounts, from which hers cannot be distinguished.

## 7. REFERENCES

[1] http://www.krebsonsecurity.com/2010/01/texas-bank-sues-customer-hit-by-800000\-cyber-heist/.

[2] http://www.schneier.com/blog/archives/2010/01/32_million_jewe.html.

[3] Symantec Internet Security Threat Report XIII. http://eval.symantec.com/mktginfo/enterprise/white_papers/b-whitepaper_internet_security_threat_report_xiii_04-2008.en-us.pdf.

[4] Symantec Report on the Underground Economy XII. `http://eval.symantec.com/mktginfo/enterprise/white_papers/ent-whitepaper_internet_security_threat_report_xii_09_2007.en-us.pdf`.

[5] A. Beautement, M.A. Sasse and M. Wonham. The Compliance Budget: Managing Security Behaviour in Organisations. *NSPW*, 2008.

[6] A. Odlyzko. Providing Security With Insecure Systems. *WiSec*, 2010.

[7] Adam Barth, Benjamin I. P. Rubinstein, Mukund Sundararajan, John C. Mitchell, Dawn Song, and Peter L. Bartlett. A Learning-Based Approach to Reactive Security. *Financial Crypto*, 2010.

[8] A. Adams and M. A. Sasse. Users Are Not the Enemy. *Commun. ACM*, 42(12), 1999.

[9] B. Laurie and R. Clayton. Proof of Work Proves not to Work. *WEIS*, 2004.

[10] M. Bishop. *Computer Security: Art and Science.* Addison Wesley, 2003.

[11] Brett Stone-Gross, Marco Cova, Lorenzo Cavallaro, Bob Gilbert, Martin Szydlowski, Richard Kemmerer, Christopher Kruegel, and Giovanni Vigna. Your Botnet is My Botnet: Analysis of a Botnet Takeover. *CCS*, 2009.

[12] C. Dwork and M. Naor. Pricing via Processing or Combatting Junk Mail. *Crypto*, 1992.

[13] C. Jackson. *Improving Browser Security Policies.* PhD thesis, Stanford University, 2009.

[14] C. Martindale. Fame is More Fickle Than Fortune: On the Distribution of Literary Fame. *Poetics*, 1995.

[15] C. Shapiro and H.R. Varian. Information Rules. *Harvard Business School Press*, 1999.

[16] W. Cheswick, S.M.Bellovin, and A. Rubin. Firewalls and Internet Security. *Addison-Wesley*, 2003.

[17] D. Liu and L.J. Camp. Proof of Work can Work. *WEIS*, 2006.

[18] D. Price. Big Science, Small Science. Columbia University Press.

[19] R. Dhamija, J. D. Tygar, and M. Hearst. Why phishing works. *CHI*, 2006.

[20] Federal Reserve Board. Survey of Consumer Finances. `http://www.federalreserve.gov/pubs/oss/oss2/scfindex.html`.

[21] D. Florêncio and C. Herley. A Large-Scale Study of Web Password Habits. *WWW 2007, Banff.*

[22] H. R. Varian. System Reliability and Free Riding. *Economics of Information Security*, 2004.

[23] C. Herley. So Long, And No Thanks for the Externalities: The Rational Rejection of Security Advice by Users. *NSPW 2009, Oxford.*

[24] C. Herley and D. Florêncio. A Profitless Endeavor: Phishing as Tragedy of the Commons. *NSPW 2008, Lake Tahoe, CA.*

[25] C. Herley and D. Florêncio. Nobody Sells Gold for the Price of Silver: Dishonesty, Uncertainty and the Underground Economy. *WEIS 2009, London.*

[26] J. Sunshine, S. Egelman, H. Almuhimedi, N. Atri and L. F. Cranor. Crying Wolf: An Empirical Study of SSL Warning Effectiveness. *Usenix Security*, 2009.

[27] K. Mitnick and W.L. Simon. *The Art of Deception: Controlling the Human Element of Security.* Wiley, 2003.

[28] C. Kanich, C. Kreibich, K. Levchenko, B. Enright, G. M. Voelker, V. Paxson, and S. Savage. Spamalytics: An empirical analysis of spam marketing conversion. In *Proceedings of the 15th ACM Conference on Computer and Communications Security*, pages 3–14, Alexandria, Virginia, USA, October 2008.

[29] L. Zhuang, F. Zhou, and J.D.Tygar. Keyboard acoustic emanations revisited. *CCS*, 2005.

[30] L.A. Gordon and M.P. Loeb. The Economics of Information Security Investment. *ACM Trans. on Information and System Security*, 2002.

[31] M. G. Kuhn. Electromagnetic eavesdropping risks of flat-panel displays. *Proc. PETS*, pages 88–107.

[32] M. Jakobsson and S. Myers. Phishing and Countermeasures: Understanding the Increasing Problem of Electronic Identity Theft.

[33] S. McClure, J. Scambray, and G. Kurtz. *Hacking Exposed.* McAfee, fifth edition, 2005.

[34] Michael Backes, Markus Duermuth, and Dominique Unruh. Compromising Reflections - or - How to Read LCD Monitors Around the Corner. *IEEE Symposium on Security and Privacy*, 2008.

[35] N.G. Mankiw. Principles of Economics. *4-th ed.*, 2007.

[36] N. Provos and T. Holz. *Virtual Honeypots.* Addison Wesley, 2007.

[37] R. Anderson. Why Information Security is Hard. In *Proc. ACSAC*, 2001.

[38] R. Anderson. Security Engineering. In *Second ed.*, 2008.

[39] R. Boehme and T. Moore. The Iterated Weakest Link: A Model of Adaptive Security Investment. *WEIS*, 2009.

[40] R. Thomas and J. Martin. The Underground Economy: Priceless. *Usenix ;login:*, 2006.

[41] S. Schechter, R. Dhamija, A. Ozment, I. Fischer. The Emperor's New Security Indicators: An evaluation of website authentication and the effect of role playing on usability studies. *IEEE Security & Privacy*, 2007.

[42] S. Saroiu, S. D. Gribble, and H. M. Levy.

Measurement and Analysis of Spyware in a University Environment. *Proc. NSDI*, 2004.

[43] S. E. Schechter, A. J. B. Brush, and S. Egelman. It's No Secret: Measuring the Security and Reliability of Authentication via "Secret" Questions. In *IEEE Symposium on Security and Privacy*, pages 375–390, 2009.

[44] E. Skoudis and L. Zeltser. *Malware: Fighting Malicious Code*. Prentice Hall, 2004.

[45] T. Moore and R. Clayton. Examining the Impact of Website Take-down on Phishing. *Proc. APWG eCrime Summit*, 2007.

[46] T.N. Jagatic, N.A. Johnson, M. Jakobsson and F. Menczer. Social Phishing. *Commun. ACM*, 2007.