Toward Econometric Models of the Security Risk from Remote Attacks

Stuart E. Schechter
Harvard University
stuart@eecs.harvard.edu

May 13-14, 2004

Abstract

Security risk regression models have been successful in estimating the likelihood of attack for simple security threats in which offensive and defensive innovations occur at a slow pace, such as burglary. Adapting regression models to the numerous and dynamic threats that face computer systems will be more challenging. One reason is that models of remote network attacks will need to account for the adversaries' ability to use the network to cover their tracks, reducing the risks of incarceration and other harm that could result from staging attacks. When our adversaries no longer face significant risk, the aggregate cost of time, effort, and other resources required to stage a successful attack is more likely to be the salient deterrent. The resource cost of attack depends on how strong the system's security is in resisting attack. I present a framework for *security risk* regression models in which regressors are chosen from four classes, one of which is this *security strength*. I then present significant refinements to previous metrics of security strength and their measurement.

1 Introducton

System security is an endeavor undertaken to reduce the risk of security breaches. In the field of risk management, risk is often defined as the expected loss due to a potential event.

$$\begin{aligned} \text{risk} &= (\text{likelihood of loss event}) \times (\text{cost of loss event}) \\ \text{security risk} &= (\text{likelihood of security breach}) \times (\text{cost of security breach}) \end{aligned}$$

For the more general case in which more than one loss event or breach may occur, *security risk* may be defined in terms of the frequency with which breaches are expected to occur (or the security breach rate.)

$$\text{security risk} = (\text{security breach rate}) \times (\text{average cost per breach})$$

To forecast security risk requires one to first forecast the rate at which different types of breaches will occur. The better the risk model, the better the security decisions that can be made using its forecasts.

Successful security risk models can be created when the circumstances that could lead to a breach are well understood, when security risk is a function of independent variables (regressors) that are directly or indirectly measurable, and when the dependence of security risk on these variables remains stationary over time. If important indicators of security risk aren't measurable or simply aren't measured, the model will suffer from omitted variable bias. Even if a model can accurately estimate past security risks using historical data, the model will be of little value if the relationships on which the model relies no longer hold. This is of particular concern in security as the dominant strategy of our adversaries is often to work outside our model and defy our predictions.

The frequency with which a system will be successfully attacked, and security breached, is likely to depend on at least four sets of factors each of which may be represented by one or more independent variables.

First, there is the question of how many individuals are in a position to attack the system. The *number of potential adversaries* is likely to be positively correlated with the frequency of security breaches and the resulting security risk. If the class of breach being studies is one caused by an insider attack, the set of potential adversaries is likely to be smaller than for an outside attack.

Secondly, there is the question of how valuable attack appears to be to your potential adversaries. The greater the *incentive to attack*, the more likely it is that a potential adversary will choose to stage an attack. Thus, the incentive to attack is also positively correlated with attack frequency and security risk.

The presence of factors that provide disincentives to attack are expected to be negatively correlated with the frequency of attacks and thus also negatively correlated with security risk. One class of disincentives is *attack risk*, or the risk of undesirable consequences to the adversary as a result of attacking the system. These consequences may include disclosure of the adversary's attack tools or techniques, reputation damage, capture and incarceration, or even physical harm.

Finally, the collective cost of equipment, effort, and other resources required to stage a successful attack also acts as a disincentive and is negatively correlated with attack frequency and security risk. The stronger the system, the greater the expected cost to successfully attack it. Thus, *security strength* is negatively correlated with attack frequency and security risk.

At the First Workshop on Economics and Information Security, I introduced economic approach for measuring the security strength of software in units of dollars [11]. These security strength metrics, which have since been significantly refined [10, 12], can be used to help forecast the more elusive security risk metrics.

In Section 2, I examine regression models of burglary that bring a level of quantitative forecasting to study of risk in home security that we in computer security have yet to achieve. I contend, in Section 3, that the key barrier to the use of regression models for forecasting the security risk from remote network

attacks has been the omission of a variable that measures security strength. In Section 4, I present significant refinements to my earlier work on measuring security strength and the theory behind them. Related work in the use of regression models in computer security is presented in Section 5, and I conclude in Section 6.

For those unfamiliar with regression models, I provide a brief introduction to their use in Appendix A.

2 Regression models in security: studies of home burglary

The problem of securing the modern home against losses due to burglary has been extensively studied. Statistical data has been collected and applied to regression analyses to estimate the likelihood of burglary for homes with different traits and different safeguards [14, 4, 3].

A sample set of independent variables (regressors), measured by Hakim et al. [4] to determine their effect on burglary rates, is shown below. Independent variables that are posed in the form of boolean (true or false) questions take the binary value of either zero, for false, or one, for true. Those not in the form of questions are scalar variables.

- Is a dead bolt on the door?
- Is an alarm installed?
- Is a car always in the driveway?
- Is a light on timer or motion sensor?
- Is a radio or television on timer?
- Are mail/papers left uncollected?
- Is home on corner of block?
- Is home $\frac{1}{4}$ mile from highway exit?
- Is home next to woods/playground?
- Value of home
- Number of children living in house
- Number of years residing in home

As the study was performed in exclusively in the city of Greenwich, Connecticut, the number of potential adversaries may not vary widely throughout the sampled population of homes. However, it is possible that those homes closer to the highway will be more likely to be targeted by a larger set of adversaries that includes thieves from large nearby cities (New York or Stamford.) The number of potential adversaries may thus effect the correlation between one of our regressors (the indicator of whether a home is a quarter mile from the highway) and the likelihood of burglary.

Burglars are more likely to find valuables and expensive homes than inexpensive ones. Thus, the incentive to attack is likely to be represented by the regressor that measures the value of the home.

Deadbolts make entering a home more difficult. The regressor that indicates whether or not the home's doors have deadbolts is thus an indicator of security strength.

Most of the remaining regressors represent factors that increase the attack risk that a burglar faces if he should decide to target the home. Cars in driveways, lights and radios on timers, and additional occupants (children) decrease a burglar's certainty that a home will be empty when he decides to enter. On the other hand, if mail is left uncollected the burglar will have increased certainty that nobody is home.

The study of Shachmurove et al. [14], using the Greenwich data, resulted in a model that could be used by homeowners in that area to make better security decisions. They showed that the deterrents that were effective against burglars were those that either increased a burglar's risk of being detected or that merely led the burglar to believe he was more likely to be detected. While alarms were most effective, lights left on timers and cars in driveways made the level of personal risk uncertain enough that a significant fraction of burglars would avoid targeting the home. More surprising was that the security measure designed to increase the difficulty of entry, the dead bolt, did not have a statistically significant effect in reducing the likelihood that a home would be burgled. Burglars preferred to risk of entering a house through a window to taking the risk of casing and approach another home in its place.

While surprising, the regression study is consistent with findings of Wright and Decker [21], who surveyed 108 burglars in the St. Louis area. They also reported that alarms were a strong deterrent, with 56 of 86 burglars reporting that they would never target a house known to have an alarm [21, page 125]. They reported anecdotal evidence that burglars are hard to deter once they've already trespassed on a property. When confronted with a dead bolts, the burglars in their study said they would break the bolt, or enter the house through a window, rather than retreating and finding a new home to break into. Windows with a single pane can be broken without making much noise and so do not greatly increase the risk of to the burglar. In the words of one of the burglars interviewed, "as long as houses are made of wood and glass, I can get 'em" [21, page 98]. Using the terminology of this paper, this statement might be translated as "no house with wood doors or glass windows is strong enough to keep me out."

If home security practices have benefited from regression models, why hasn't the computer industry applied similar models to measure security risk as a function of system architecture and safeguard choice? On reason is that computer systems are far more complex and heterogenous than homes. Homes are built of "wood and glass" and serve a common function. Software may be written using different algorithms coded in different languages. Because different software packages perform dramatically different functions for their users, their architectures are varied, as are the classes of attack to which they are vulnerable.

Perhaps most importantly, many computer crimes can be committed from a safe distance, making risk far less of a deterrent for these computer criminals than it is for burglars.

3 Remote attacks and the role of security strength

Remote attacks via networks are not likely be deterred by the introduction of safeguards that increase the risk that an attacker will be detected. Global networks, such as the Internet, enable criminals to stage attacks with little risk of retribution. Whereas it was once the case that criminals could flee to foreign countries after a crime, a network enables an adversary to seek refuge first and then to attack from a safe distance. Foreign jurisdictions that do not consider the attack to be a crime, or that do not have the interest or resources to help identify and extradite the perpetrator, are ideal for such purposes.

Staging an attack without being identified is preferable to staging an attack and avoiding prosecution. Network attackers regularly hide their identities by routing their communicating through a sequence of distant systems. The longer the trail, the harder it is trace its source. If each step in the route lies in a different jurisdiction, and the trail can only be unravelled one step at a time, tracking the adversary will require gaining the cooperation of each jurisdiction. If enough steps lie between the adversary and the target, tracking the source of the attack is likely to take far more time than the attack itself. Unless meticulous logs are kept by all compromised parties, the trail will disappear before it can be traced.

A system's defenses can do little to increase the risk that a remote attacker will be identified, let alone punished. When an adversary can attack with impunity, the time, effort, and other resources required to stage an attack become significant factors in the adversary's choice to target a system. The stronger the system, the more resources are likely to be believed to be required.

The decline in effectiveness of strategies that rely on detecting and punishing the adversary, leaving vulnerabilities unguarded and systems weak, is both inevitable and already well underway. Systems with unpatched vulnerabilities, for which exploits have been written, are now compromised within minutes of being exposed to the Internet [20]. The steady increase in network attacks seen in recent years was inevitable given the decline in both the risk and the cost of attack. The automation of tools such as attack scripts have further reduced the cost to attack systems. The steady increase in the number of vulnerabilities discovered, and left unpatched, has ensured a steady supply of attack scripts. As home users have acquired high speed, always-on, Internet access, their poorly guarded machines have become available for criminals to route traffic through, making it even easier for these criminals to mitigate their risk.

Others, including Beattie et al. [1], have analyzed the consequences of leaving a vulnerability exposed to avoid the costs and risk of patching. However, such work is predicated on statistical relationships that have not stood still. Their work has become quickly outdated as the risk of leaving vulnerabilities unpatched has rapidly increased. One reason has been the increase in self-reproducing malware, such as the the worms predicted by Staniford et al. [17] and later demonstrated by the Slammer worm as documented by Moore et al. [9]. What's more, worm development kits promise to reduce the time and skill required to create worms that exploit newly discovered vulnerabilities. Hybrid

pathogens now bypass organizations' firewalls by travelling through email like a virus (which, unlike worms, require human interaction to spread), then attack organizations from the inside exploiting vulnerabilities to once again spread like a worm (without human intervention).

Past attempts to bring the quantitative approaches of insurance and risk management to the measurement of security risk have failed because, unlike earthquakes or burglars, the threats posed to computer systems are far from stationary.

On the other hand, it may still be possible to model the likelihood of attack on a system as a function of the number of potential attackers, how lucrative the attack is perceived to be by those potential attackers, how much risk one must be willing to accept to attack the system, and how strong the system is perceived to be. The measurement of the security strength of software systems is thus essential if we are to build working regression models to predict their effectiveness in reducing risk, especially if that strategy that includes efforts to make attack more difficult.

4 Measuring security strength

For any given threat, a system is only as strong as it is difficult for an adversary to succeed in attacking it. Safeguards, such as cryptosystems, that have been used to prevent certain classes of attacks, have long been measured by how difficult they are to circumvent. A security strength metric attempts to quantify the time, effort, and other resources that are expected to be required to bypass a system's safeguards to carry out an attack. Because security strength metrics gauge security from the perspective of the adversary, they complement security risk metrics, which measure security from the perspective of those the safeguards are intended to defend.

Because security strength measures the resources needed to breach a system's security, it is a fundamentally economic metric. Formal methods of Computer Science, such as information theory, have been used to address questions of security strength when these questions can be translated into assertions of what can and cannot be calculated. For example, Shannon used information theory proved the one-time pad cipher to be secure against certain classes of attack [16] regardless of the resources available to the attacker. Shamir used similar techniques to prove security properties of his secret sharing scheme [15]. However, the applicability of computational approaches to security strength is extremely limited. Even today's best public key cryptosystems are only as strong as certain algorithmic problems, on which their security is predicated, are difficult to solve. Software running on networked computers, which may be attacked from a remote location, is too complex to be addressed using formal proofs of computability.

While the complexity of software hinders security strength measurement, software has other properties that favor measurement. The ease with which verifiably exact replicas can be made ensures that the cost of measurement can

be amortized over each copy. Identical copies can be tested for vulnerabilities in parallel. The ease with which digital data can be compared enables end users to verify that the software they are running is in fact the software that was demonstrated to have a measured security strength.

The practice of gauging the strength of a software system has long revolved around the search for vulnerabilities. The quantity and wide availability of software tools that exploit vulnerabilities to attack systems, which often appear very soon after vulnerabilities are publicized, ensure that the strength of systems with known vulnerabilities is too small to be worth measuring. It is only interesting to estimate the security strength of systems in which there are no known vulnerabilities or those in which vulnerabilities are known but require considerable resources to exploit.

In order to breach the security of a system with no known vulnerabilities, a new vulnerability must be found and a technique to exploit the vulnerability to breach security (known simply as an *exploit*) must be developed. The strength of the system is thus dominated by the cost to find a vulnerability and create an exploit. Since primitive exploits are often required in order to demonstrate the existence of a vulnerability, these costs may be considered together as the cost of finding a vulnerability[1].

One of the difficulties in measuring security strength is that it is a function of your adversary's costs, not your own. It's impossible to research the skill of every potential adversary, the value they place on their time, and the other resource costs they must expend to find a vulnerability in a software program. What's more, the adversaries themselves are not likely to not know how much time and resources they will need to expend to accomplish their goals.

To understand why even the adversary may not be able to measure his own costs, assume that finding a vulnerability is an essential step in breaching the security of a system. There are a series of tasks the adversary can perform to look for vulnerabilities, from inspecting code to writing and executing tests. To maximize his productivity, the adversary will start with the tasks that have the greatest chance of success in finding a vulnerability per unit cost. Diminishing expected returns result because the tasks with the highest expected profitability are executed first. We can see in Figure 1 that individuals perceive this cost not as a single value, but as a cumulative probability distribution. The chance of success in finding a vulnerability increases with total investment (the first derivative is positive), but the chance of success for each additional dollar invested is smaller than for the previous dollar (the second derivative is negative).

An economically rational individual will only perform tasks so long as their expected return is greater than their expected cost. If the individual believes a

[1] It is possible for the cost of exploiting a vulnerability to exceed the cost of discovering it. An extreme example of such a vulnerability is a back-door guarded by public key cryptography. That is, a vulnerability intentionally placed in the code in order to grant access to anyone who knows a secret key. While such a vulnerability may be easy to discover, the secret key may be impossible to find as it need not be included in the code itself. Writing an exploit without knowledge of the secret key requires one to accomplish the monumental task of breaking the public key cryptosystem to find the secret key.

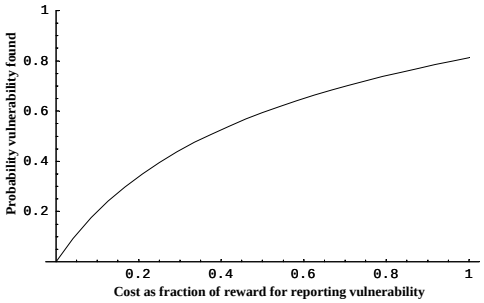

Figure 1: A cumulative probability distribution that represents an individual's beliefs of the likelihood that he or she can find a vulnerability in a system (the y axis) for the cost represented by the x axis.

vulnerability is worth r dollars, the cost of a task i is c_i, and the probability that task i will result in the discovery of a vulnerability is p_i, then he will perform the task only when $c_i \leq p_i \cdot r$, or equivalently when $\frac{c_i}{p_i} \leq r$. In fact, while a risk neutral individual will continue to perform tasks when $\frac{c_i}{p_i} < r$ and be indifferent to performing tasks when $\frac{c_i}{p_i} = r$, a risk averse adversary will not search for vulnerabilities when $\frac{c_i}{p_i} > r - \epsilon$ for some positive measure of risk aversion ϵ.

There is no reason to believe that, given this uncertainty, an individual will expend the same amount of resources to find a vulnerability as he would if he knew the cost beforehand. Perceptions may be updated between tasks, as c_i and p_i may not need to be calculated until task $i - 1$ is complete. However, updating these estimates cannot provide a guarantee that they will be accurate. It is the perceived expectation of the cost, not the true cost, that determines how many resources an adversary will be willing to spend to find vulnerabilities in a system. Thus, a metric of security strength that incorporates perceptions of cost may not only be as valid as a measure that includes only true costs, it may be more so[2]. While in my earlier work I proposed that security strength be approximated by a metric I referred to as the *cost-to-break* [11], it is more accurate to call this metric perceived cost-to-break, which can be measured by determining the market price one would have to pay to acquire a vulnerability.

An adversary intent on breaching the security of a computer system may search for a vulnerability himself or he may exploit the labor market to pay someone else to do it. If we assume the worst, the individual with the lowest

[2] Security is not the only field in which perception can often trump reality. The theory of investment tells us the true value of a firm is the net present value of the income stream the firm will return to its investors. However, a firm with a higher perceived value will be more likely to obtain debt financing, make deals, and retain employees. Such a company will thus be more likely to succeed, and return money to its investors, than a firm that is identical except for perceptions. This is why perceptions and true value are inseparable. What's more, the true value of a firm can never be known so long as it is a going concern with an uncertain future. The perceived value of a publicly traded firm is usually readily available and falls within a tight range between the market bid and ask price.

cost of finding and exploiting a vulnerability and the individual with the most to gain from this knowledge are one and the same. Equivalently, they may be different individuals who reach an agreement by which the vulnerability information is sold.

The market price to find and report a vulnerability, or *MPV* estimates the cost of finding a vulnerability in a software system as measured by a market in which anyone, including those we might consider to be our adversaries, can participate. The MPV bid price is the highest price offered to buy a previously unreported vulnerability. The MPV ask price is the lowest price at which a seller is asking to report a vulnerability. Transactions occur when the bid price and the ask price meet. The MPV at all other times falls within the range of the bid price (the lower bound) and the ask price (the upper bound).

The reward offered by a firm for vulnerabilities in its product, which constitutes the bid price for the reporting of a vulnerability, can be used as a metric of security strength in models that forecast security risk. However, this value is a lower bound and may differ from the market perceptions of the true price one would have to pay to obtain the next vulnerability. We would also like to know how long one would have to wait to obtain a vulnerability after offering a given price.

To more accurately gauge these values, regression models could be built using data generated by vulnerability reporting markets. One could estimate MPV by modelling it as a function of the current bid price, the price at which previous vulnerabilities were reported and the amount of time that has passed since each was reported, and any other factors of import. The same approach could be used to forecast the probability that a new vulnerability will be reported for a given price within a specified period of time.

5 Prior regression models in computer security

While we have yet to see software security regression studies to forecast security risk, as we have seen for home burglary, a limited number of studies have used statistical analysis and regression models of security data for other goals.

One of the earliest applications of statistical methods to security risk was a 1990 study by Detmar Straub [19], who showed that organizational commitment to invest in computer security has a statistically significant effect. Because Straub's intent was to show that certain actions have effect, rather than to measure the amount of risk reduction, he did not use a full regression model. Because the study was performed well before most corporations were connected to the Internet, the results are primarily of interest when studying scenarios involving local attack. Another key limitation of this study is the poor resolution of the independent variables. For example, one independent variable represents the number of security programs installed on a system, without regard to what these programs do or how up to date they are.

Straub's ability to show statistically significant effects using such imprecise independent variables would imply that regression models using more precise

measures could have promise. Indeed, regression models have been used to measure the rate of security incidents that exploit a vulnerability, as a function of the time since the vulnerability was discovered [2]. Regression analysis has also been used to show that the overall frequency of security incidents is increasing [5]. Given the potential of regression models for measuring security risk, it is not surprising that such studies have been proposed to study computer fraud [6] and insider attacks [13]. However, the resulting studies have not been forthcoming.

While not directly estimating security risk, Moitra and Konda [7, 8] used regression models to generate the constants for a stochastic model of yet another related metric, system survivability. The problem with this study was not the regression model, but the stochastic model. The stochastic model assumed attacks arrived at random, and so the utility of this work is limited by the underlying assumption that attackers behave in a random manner.

6 Conclusion

Regression models have proven themselves valuable for estimating the effectiveness of safeguards in reducing the security risks posed by well understood threats. Specifically, they have been successful when security risk has been a function of the presence of safeguards that cause adversaries to perceive unacceptable risk that an attack will result in consequences they find undesirable. Successful models have also required that the factors affecting the adversaries' choices be measurable and stationary, as one would expect for threats where the rate of innovation is low, such as burglary.

Given the impunity with which remote criminals can probe our software and networks, the security of our systems against remote attacks rests upon how difficult we can make it for our adversaries to carry out these attacks. This is a question of security strength. Given that the vulnerabilities and safeguards in networked software systems are themselves not stationary, historical security risk data that is not understood in the context of security strength is of little value for forecasting future risk.

Unfortunately, today's software products are released without any measure of their strength. Until they are, we will not be able to create accurate security risk models for remote network attacks.

A A brief introduction to regression models

Regression models are tools used to estimate a measure, or *dependent variable*, as a function of a set of other measures, or *independent variables*[3]. For example, given a dependent variable Y that appears to change linearly with the value of independent variables X_1 and X_2, we might estimate Y with a value $\hat{Y}$, that is defined via a function f.

$$\hat{Y} = f(X_1, X_2) = \beta_0 + \beta_1 X_1 + \beta_2 X_2$$

For a given data point i, the difference between the estimate $\hat{Y}_i$ and the actual value of Y_i that was measured is the error term u_i.

$$Y_i = \hat{Y}_i + u_i = f(X_{1,i}, X_{2,i}) + u_i = \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + u_i$$

A regression is the process by which values are assigned to the β constants so that $\hat{Y}$ will best estimate Y. Most linear regressions select β values so as to minimize the sum of the squares of the error terms, $\sum_{\forall i} u_i^2$. For more complex functions, the maximum likelihood estimator is commonly used. Further error estimates are used to measure how well the model fits the data. One may also calculate the likelihood that a relationship between the dependent variable and one or more of the independent variables is statistically significant. To model more complex relationships, regressions can be performed on functions with non-linear terms, and even with terms in which variables interact (e.g. $\hat{Y} = \phi\left(\beta X_1 \sqrt{X_2}\right)$).

Once the regression has selected the optimal β values, we can assign those values as constants and use $f(\vec{X})$ as a functional model for estimating the effect on Y of changing one or more independent variables. If a regression model is to accurately forecast the future value of its dependent variable, the underlying relationships between this variable and the independent variables, on which the model depends, must remain stationary over time.

[3] For a more detailed introduction to regression analysis, I recommend Stock and Watson's textbook on Econometrics [18].

References

[1] Steve Beattie, Seth Arnold, Crispin Cowan, Perry Wagle, Chris Wright, and Adam Shostack. Timing the application of security patches for optimal uptime. In *Proceedings of The 16th USENIX Systems Administration Conference (LISA 2002)*, November 3–8, 2002.

[2] Hilary K. Browne, William A. Arbaugh, John McHugh, and William L. Fithen. A trend analysis of exploitations. In *Proceedings of the IEEE Symposium on Security and Privacy*, pages 214–229, May 2001.

[3] Tracey Budd. Burglary of domestic dwellings: Findings from the british crime survey. Technical report, United Kingdom Home Office Crime Reduction Programme Unit, April 1999.

[4] Simon Hakim, George F. Rengert, and Yochanan Shachmurove. Knowing your odds: Home burglary and the odds ratio. Technical Report CARESS Working Paper 00-14, University of Pennsylvania Center for Analytic Research in Economics and the Social Sciences, September 2000.

[5] John D. Howard. *An Analysis Of Security Incidents On The Internet: 1989 - 1995*. PhD thesis, Carnegie Mellon University, April 7, 1997.

[6] Lindsay C. J. Mercer. Fraud detection via regression analysis. *Computers & Security*, 9(4), June 1990.

[7] Soumyo D. Moitra and Suresh L. Konda. A simulation model for managing survivability of networked information systems. Technical Report ESC-TR-2000-020, Carnegie Mellon Software Engineering Institute, December 2000.

[8] Sournyo D. Moitra and Suresh L. Konda. The survivability of network systems: An empirical analysis. Technical Report CMU/SEI-2000-TR-021, Carnegie Mellon Software Engineering Institute, December 2000.

[9] David Moore, Vern Paxson, Stefan Savage, Colleen Shannon, Stuart Staniford, and Nicholas Weaver. Inside the Slammer worm. *IEEE Security and Privacy*, 1:33–39, July 2003.

[10] Stuart E. Schechter. How to buy better testing: Using competition to get the most security and robustness for your dollar. In George Davida, Yair Frankel, and Owen Rees, editors, *Proceedings of Infrastructure Security International Conference (InfraSec 2002)*, pages 73–87. Springer-Verlag, LNCS 2437, October 1-3, 2002.

[11] Stuart E. Schechter. Quantitatively differentiating system security. In *The First Workshop on Economics and Information Security*, May 16-17, 2002.

[12] Stuart E. Schechter. *Computer Security Strength and Risk: A Quantitative Approach*. PhD thesis, Harvard University DEAS, June 2004.

[13] E. Eugene Schultz. A framework for understanding and predicting insider attacks. *Computers & Security*, 21(6), October 2002.

[14] Yochanan Shachmurove, Gideon Fishman, and Simon Hakim. The burglar as a rational economic agent. Technical Report CARESS Working Paper 97-07, University of Pennsylvania Center for Analytic Research in Economics and the Social Sciences, June 1997.

[15] Adi Shamir. How to share a secret. *Communications of the ACM*, 22(11):612–613, November 1979.

[16] Claude Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:398–403, 1948.

[17] Stuart Staniford, Vern Paxson, and Nicholas Weaver. How to 0wn the Internet in your spare time. In *Proceedings of the 11th USENIX Security Symposium*, August 7–9, 2002.

[18] James H Stock and Mark W. Watson. *Introduction to Econometrics*. Pearson Education, Boston, MA, 2003.

[19] Detmar W. Straub. Effective IS security: An empirical study. *Information Systems Research*, 1(3):255–276, January 16, 1990.

[20] The Honeynet Project. *Know Your Enemy: Revealing the Security Tools, Tactics, and Motives of the Blackhat Community*. Addison-Wesley, 2001.

[21] Richard T. Wright and Scott H. Decker. *Burglars on the Job: Streetlife and Residential Break-ins*. Northeastern University Press, Boston, 1994.