A Framework for Classifying and Comparing Models of Cyber Security Investment to Support Policy and Decision-Making[1]

Rachel Rue, Shari Lawrence Pfleeger and David Ortiz

RAND Corporation

1200 South Hayes Street

Arlington, Virginia 22202-5050

Rachel_Rue@rand.org, Pfleeger@rand.org, David_Ortiz@rand.org

Abstract

The threat to cyber security is real and growing. Organizations of all kinds must take protective measures, but effective resource allocation is difficult. This situation is due in part to uncertainty about the nature and severity of threats and vulnerabilities, as well as about the effectiveness of mitigating measures. A variety of models have been proposed to aid decision makers. We describe a framework to analyze and compare models, and illustrate our framework with an analysis of three commonly-used types of models.

Introduction

Deciding how best to invest resources in cyber security is not straightforward. The difficulty is compounded by multiple uncertainties about threats and vulnerabilities, about the consequences of a successful attack, and about the effectiveness of mitigation measures. The sources of uncertainty range from the shifting uses of information technology to the evolving nature of the threats. Moreover, the consequences of not making good decisions about appropriate investment in cyber security resources become more severe as organizations store more and more types of information of increasing sensitivity and value. Methods of accessing the information are expanding to include a greater number of mobile and remote devices. More methods of access to information translate into at least two situations of concern: more modes of attack and an increased probability that an attack will be successful. Moreover, mitigating the threats by understanding the motives and goals of attackers requires cultural and political expertise that often does not reside within organizations.

[1] This work was supported by the Economics of Cyber Security project of the Institute for Information Infrastructure Protection (I3P) under award number **2003-TK-TX-0003** from the Office for Domestic Preparedness/Office of Justice Programs and the Department of Homeland Security. The presentation is based on RAND Corporation research and authors' opinions. Parts of the presentation describe work in progress that has not undergone RAND quality assurance procedures.

Given the challenge of ensuring cyber security under conditions of uncertainty, how can organizations determine appropriate measures to enhance cyber security and allocate resources most effectively? Models and model-based tools exist to assist in this decision-making, but it is essential to understand which models are most appropriate for which kinds of decision support. This paper explores the attributes of economic models of cyber security, provides a framework for evaluating whether a model is appropriate for a particular application, and illustrates the use of the framework by discussing in detail how several types of commonly-used models can be assessed and compared. The purpose of the assessment and comparison is to ensure that decision-makers use the best models for the job at hand, and to help decision-makers understand the strengths and weaknesses of each modeling technique.

Many models have been proposed to help decision makers allocate resources to cyber security, each taking a different approach to the same fundamental question. Macro-economic input/output models have been proposed to evaluate the sensitivity of the U.S. economy to cyber-attacks in particular sectors (Santos and Haimes 2004) and the potential for underinvestment in cyber security (Garcia and Horowitz 2006). More traditional econometric techniques have been used to analyze the loss of market capitalization after a cyber-security incident (Campbell et al. 2003). Methods derived from financial markets have been adapted to determine the "return on security investment" (Geer 2001; Gordon and Loeb 2005; Willemson 2006). Case studies of firms have been performed to characterize real-world decision making with respect to cyber security (Dynes, Brechbuhl, and Johnson 2005; Johnson and Goetz 2007; Pfleeger, Libicki and Webber 2007). Heuristic models rank costs, benefits, and risks of strategies for allocating resources to improve cyber security (Gal-Or and Ghose 2005; Gordon, Loeb, and Sohail 2003). Because investing in cyber security is an exercise in risk management, many researchers have attempted to characterize behavior through a risk management and insurance framework (Baer 2003; Conrad 2005; Farahmand et al. 2005; Geer 2004; Gordon, Loeb, and Sohail 2003; Haimes and Chittester 2005; Soo Hoo 2000; Baer and Parkinson 2007). Recognizing that potential attackers and firms are natural adversaries, researchers have also applied methods from game theory, and developed real games, to analyze resource allocation in cyber security (Gal-Or and Ghose 2005; Horowitz and Garcia 2005; Irvine and Thompson; Irvine, Thompson, and Allen 2005).

Each model is based on a different set of assumptions regarding:

- The characteristics of information systems,
- The motivations of organizations to protect information,
- The goals of attackers, and
- The data required for validation of the model.

No single model by itself can provide a comprehensive approach to guide investments in cyber security. Indeed, it is often unclear how a particular model for cyber security can be used in practice, using actual instead of theoretical data to support corporate or organizational decision makers. Rather than expecting a decision maker to rely on a single, comprehensive model, we propose that decision makers and their organizations understand how to evaluate and use several models in concert, either to triangulate and

find an acceptable strategy for investing in cyber security, or to address multiple aspects of a larger problem.

The framework we describe below can be used for assessing and comparing the value of different models in light of these several needs. Our framework is inspired by and extends two approaches used successfully in other venues to evaluate the appropriateness of decision support models: Morgan and Henrion's (1990) framework for quantifying uncertainty in policy-based economic models, and an accounting framework previously used to provide guiding principles for formulating and evaluating policies affecting greenhouse gas emissions (The GHG Protocol for Project Accounting 2005).[2]

The remainder of the paper is organized in three sections. The first section describes the framework for comparing economic models of cyber security. The second section illustrates the framework's utility by applying it to three commonly-used cyber security economic models. The third section concludes with observations on broader application of the framework.

Approaches to Modeling Cyber Security for Policy

Classifying Models

This section provides a framework for classifying and comparing economic models of cyber security. A model is an abstraction of real world phenomena. In its simplest form, a model transforms inputs to outputs via a mathematical or logical relationship. For example, Hooke's Law states that the opposing force of a spring (output) is proportional to the displacement of the spring from equilibrium (input). The mathematical relation simplifies the complex physical phenomenon relating stress and strain to a single equation, and is valid within a margin of error for a range of displacements.

The application of a model like Hooke's Law is fairly straightforward, in part because there are few variables, and in part because variable values are easily measured. Because economic models attempt to characterize human decision-making, they tend to be far more complex. They necessarily make several kinds of assumptions about the human context.. In the following paragraphs we describe the aspects of an economic model and its assumptions that we believe most significantly affect how and where it can be successfully applied.

The type or form of a model is its mathematical structure and overall approach. The structure determines what kind of inputs are needed, how computationally complex it is, whether it is deterministic or stochastic, and so on.. The overall approach is reflected in the choice of features and relationships, and in the way the model is applied. That is, we can glean the approach by looking at which features of the world are represented as essential, and whether the model is meant to be used (for example) to calculate exact outputs, to compare features of different scenarios, or to explore what happens when parameters are varied.

[2] We are grateful to Jeffrey Hunker at the Heinz School of Public Policy at Carnegie Mellon University for suggesting the GHG framework.

The model's intended use determines the assumptions to be made about the motivation and goals of the decision-maker. Some models are aimed at the firm, which may be contemplating (for example) the purchase of cyber-insurance; others are aimed at policy-makers, who are attempting to deploy limited resources to combat threats to the information infrastructure. But applying even a well-defined model at the enterprise level can be difficult because within a firm there may be different and conflicting goals, and different estimates of costs and benefits. Decision makers within organizations have heterogeneous perceptions of threats and risks. For example, departments specializing in information technology often think in terms of preventing, detecting, and responding to specific types of attacks. However, they often neglect the challenge of resilience in the face of attacks and information recovery after successful attacks; it is a difficult management, legal, and customer service challenge to determine the best strategies for maintaining operations when critical information is stolen, corrupted, inaccessible, or destroyed.

Assumptions are also made about the inputs and parameters used in the model. They are sometimes not well understood, difficult to quantify, or both, so simplifying assumptions are made about the mathematical form and values of relevant inputs and parameters. Most models have a set of parameters that need to be estimated before they can be applied; for example, to calculate the value of a financial option, one must know the volatility of the underlying asset and the risk free-rate of return. To illustrate the importance of these assumptions, consider that stock options and derivative financial instruments are priced based on the presumed behavior of an underlying asset, typically a stock or commodity (Hull 1997). "Real" options propose using the same analytical methods for different assets, typically those not traded on an exchange. The assumptions regarding the behavior of a stock over time, which hold true only under certain circumstances in financial markets, might not apply to the new asset in a "real" options framework, a difference that the builder of the model, and the policy maker taking its advice, need to consider.

In addition, a model makes assumptions to simplify phenomena and to focus attention on critical behaviors: Leontieff models assume that economic outputs are related linearly to economic inputs; this assumption allows more detailed study of the relationships among these factors, but only for small relative changes in their values. The assumption of linearity is necessary to make the model computationally tractable, but it limits the economic scope within which the model is valid. Most models require simplifying assumptions about the mathematical form of functions used in the model; these assumptions limit the domain of applicability of a model. For instance, Leontieff models are applicable where changes in input values are relatively small; similarly, linear models of springs are valid only for a specified range of displacements.

An additional difficulty in choosing an appropriate model for a given type of decision is that often the relevant data are not available. Models are useful only when there are valid and appropriate datasets to inform them. Historical data are often needed to show that a given type of model, with all of its simplifying assumptions, has in fact been useful in the past, and under what conditions it has been useful. Highlighting the data required to validate the use of a model can assist researchers in understanding which data sets should be solicited with surveys, interviews and automated tools.

Together, the assumptions made by a model, the data needed to support it, and its domain of applicability determine the types of decisions that the model supports, and the conditions under which the model may be applied to other situations. Thus, when deciding which model(s) to use, we want to explore the characteristics that show their purpose, application, requirements for data, and sources of uncertainty. By modifying the approach of Morgan and Henrion (Morgan and Henrion 1990), we have built Table 1, below, to list characteristics that will be helpful in classifying models of cyber security economics.

Table 1: List of characteristics that are used to describe cyber security economic models.

Characteristic	*Description*
Type or form	The class of model and its mathematical structure
History and previous applications	When and for what purpose the model was originally developed and where it has been applied successfully
Underlying assumptions	Includes simplifications made to enable easier application
Decisions that the model supports	The types of decisions that a decision-maker would be able to substantiate through proper application of the model
Inputs and outputs	The quantities or attributes that the model manipulates
Parameters and variables	Elements that affect the way in which the model transforms inputs to outputs
Applicable domain and range	Temporal and physical ranges of inputs, outputs, parameters, and variables that the model describes
Supporting data	Evidence that the model accurately represents the phenomena of interest

Comparing Diverse Models of Investment in Cyber Security

The entries in Table 1 characterize a given model, and can be used to compare models with each other, particularly for suitability for a given task. In addition we have found it useful to articulate a set of guiding principles, expressed as questions about each model, to be applied in evaluating and comparing models, as well as in developing and making use of them. These principles are suggested by a methodology used to compare different projects in terms of greenhouse gas (GHG) emissions reduction (The GHG Protocol for Project Accounting 2005). Although the GHG protocol may seem a strange choice, there are in fact underlying similarities. We know that cyber attacks have adverse economic effects, and that specific compelling examples exist to suggest particular actions in very particular circumstances. But the complete nature of the vulnerabilities, threats, and risks

to a system is uncertain. In the same way, greenhouse gases involve vulnerabilities, threats and risks that require a system-wide analysis. In both cases, comparing alternatives requires a consistent and transparent methodology. The goals of a cyber security economic comparison are:

- To enhance the credibility of economic models of cyber security by applying common accounting concepts, procedures, and principles, and
- To provide a platform for harmonizing different project-based modeling initiatives and data collection programs.

The *baseline scenario* is the canonical set of inputs, outputs, parameters, and variables that a model describes. The baseline scenario is commonly referred to as the "business as usual" case and is the one in which no action is taken by decision makers. Changes to inputs, values, and parameters represent (depending on the model) actions, investments in cyber security, emerging threats and vulnerabilities, or cyber security events. The change in the outputs from the baseline scenario illustrates to the decision maker the value of one course of action over another.

The principles described below also enable us to compare the *forms of the outputs*. All outputs have common temporal and quantitative characteristics. For example, the outputs of game theoretic models are strategies, and the outputs of insurance-valuation models are probabilistic descriptions of returns. By comparing the change in outputs from the baseline scenario, we can assess the performance of particular policies. The fidelity of the output to existing data and the relevance to actual decisions are essential. A key purpose of comparing models is to put them in a real-world context. The questions below enable us to contrast one model with another along several dimensions, each of which emphasizes the model's appropriateness for its intended use. Thus, the questions highlight the significance of model characteristics; they also helps to reveal gaps between models and the scenarios in which they are intended to be used. Making more explicit the strengths and weaknesses of each model, the model evaluation and comparison enable model developers and model users to understand the best ways to assemble needed data, run models, and present output and conclusions.

- **Is the model relevant?** *Does the model use data, methods, criteria, and assumptions that are appropriate for the intended use of reported information?* The quantification of inputs and outputs should include only information that users (of the models and of the results) need for their decision-making. Data, methods, criteria, and assumptions that can mislead or that do not conform to carefully defined model requirements are not relevant and should not be included.
- **Is the model complete**? *Does the model consider all relevant information that may affect the accounting and quantification of model inputs and outputs, and complete all requirements?* All possible effects should be considered and assessed, all relevant technologies or practices should be considered as baseline candidates, and all relevant baseline candidates should be considered when building and exercising models. The model's documentation should also specify how all data relevant to quantifying model inputs should be collected.

- **Is the model consistent?** *Does the model use data, methods, criteria, and assumptions that allow meaningful and valid comparisons?* The development and use of credible models requires that methods and procedures are always applied to a model and its components in the same manner, that the same criteria and assumptions are used to evaluate significance and relevance, and that any data collected and reported will be compatible enough to allow meaningful comparisons over time.
- **Is the model transparent?** *Does the model provide clear and sufficient information for reviewers to assess the credibility and reliability of a model and the claims derived from it?* Transparency is critical, particularly given the flexibility and policy-relevance of many decisions based on the models' outputs. Information about the model and its usage should be compiled, analyzed and documented clearly and coherently so that reviewers may evaluate its credibility. Specific exclusions or inclusions should be clearly identified, assumptions should be explained, and appropriate references should be provided for both data and assumptions. Information relating to the model's "system boundary" (i.e. the part of the problem addressed by the model)[3], the identification of baseline candidates, and the estimation of baseline data values should be sufficient to enable reviewers to understand how all conclusions were reached. A transparent report will provide a clear understanding of all assessments supporting quantification and conclusions. This analysis should be supported by comprehensive documentation of any underlying evidence to confirm and substantiate the data, methods, criteria, and assumptions used.
- **Is the model accurate?** *Does the model reduce uncertainties as much as is practical?* Uncertainties with respect to measurements, estimates, or calculations should be reduced as much as is practical, and measurement and estimation methods should avoid bias. Acceptable levels of uncertainty will depend on the objectives of the model and the intended use of the results. Greater accuracy will generally ensure greater credibility for any model-based claim. Where accuracy is sacrificed, data and estimates used to quantify a model's inputs should be conservative.
- **Is the model conservative?** *Does the model use conservative assumptions, values, and procedures when uncertainty is high?* The impact of a model should not be overestimated. Where data and assumptions are uncertain and where the cost of measures to reduce uncertainty is not worth the increase in accuracy, conservative values and assumptions should be used. Conservative values and assumptions are those that are more likely to underestimate than overestimate changes from the baseline or initial situation.

We add an additional criterion to the GHG Protocols:

[3] The system boundary allows the reader to understand each activity included in the model, and the inputs and outputs associated with each activity. That is, it defines the scope of the model, enabling the reader to determine what is included in the model and what is excluded from the model's consideration.

- **Does the model provide insight?** *Does the model state clearly the nature of the insights that are provided by the model?* Models may in some cases not serve to generate a specific result, but rather to provide a means for decision makers to better understand and gain additional useful insights into the complex problems they face. Thus, the fact the answer is '2' is far less important in some cases than that the model offers additional understanding of complex interactions.

Applying the Framework to Economic Models

Models of a wide variety of types have been constructed to represent various aspects of the economics of information security. For example, the 2006 Workshop on the Economics of Information Security included presentations using beta binomial models, one-factor latent risk models, multivariate regression models, and a two-stage non-cooperative Cournot game. To illustrate the utility of the framework presented in the previous section, we have chosen three specific model types to analyze and explore:

- *An accounting model*. Gordon and Loeb's application of accounting principles to cyber security economics to determine optimal investments in cyber security. (Gordon and Loeb, 2002) Its output is the marginal rate of return on security investment. Based on assumptions about the form of the security function, Gordon and Loeb conclude that, in many natural situations, it makes economic sense to invest only a fraction of the value of an information asset in controls to protect it.
- *A game-theoretic model*. Varian's game theoretic model to explore situations in which a system is used by many individuals, but individuals make self-interested choices about how much effort to expend in order to keep the system running. (Varian, 2004) If each individual organization commits resources only to maximize its own net benefit, the resulting distribution of costs and benefits may deviate from what is socially optimal. In addition to demonstrating that this result occurs in a number of natural cases, he uses the model to evaluate several proposed policy changes that change individual cost functions so as to make the amounts of individually optimal investments equal to what they would be in the socially optimal case.
- *An input/output model.* Andrijcic and Horowitz's input/output model to estimate the macroeconomic effects of intellectual property theft on the U.S. economy. (Andrijcic and Horowitz, 2004) It combines a model of probable foreign sources of intellectual property theft, an equity model, and an input/output model of the effect of terrorism on the U.S. national economy developed by Santos and Haimes. (Santos and Haimes, 2004). The model is applied to data from the U.S. Bureau of Economic Analysis.

In what follows, we analyze and compare each of the three models using the framework described in the previous section. The analysis addresses what is omitted from the models, the difficulty in estimating inputs, and the assumptions that are not likely to be met in the real world. This analysis is meant not as a criticism of individual research papers but rather as examples of how the research must be enhanced before the models will be ready for practical use by corporate executives. When we focus on simplifying assumptions, for example, the purpose is not to object to the assumptions, but rather to

make clear the potential danger of applying a model without understanding how accurately an assumption approximates the decision maker's real world context, and how much the outcome depends on its accuracy.

To enable ease of comparison, we begin our analysis by clarifying our terminology Both the accounting model and the game-theoretic model define a function that takes security expenditures as input and produces increased security levels as output. This type of function occurs regularly in economic models of cyber security. We call any such function a *security function.*

Accounting Model

The inputs to the accounting model are:

- A division of all information controlled by an organization into disjoint information sets,
- For each information set, an estimate of its value (i.e., the cost to the organization if it is damaged, stolen, or otherwise abused),
- For each information set, an estimate of its vulnerability, and
- The mathematical form of the security function.

Discussion of inputs:

Disjoint information sets. The input to the model is actually a single information set. This presupposes that an organization has been able to divide its information into disjoint sets in a way appropriate to the model. Guidance is required as to what criteria should be used to define the sets. Should a set be defined by the value of the pieces of information in it? (Gordon and Loeb suggest that it may make sense to divide information into low, medium, and high value sets.) Should it be defined in terms of connectedness or access? Should a set be defined in terms of the threats to which it is susceptible? Which of these ways allows the model to work best? Are there some threats for which the model is invalid? The answers depend both on the mathematical structure of the model and on empirical facts about the way attacks are targeted and spread through information systems. Both the structure and facts must be well-understood for the model to be used appropriately.

Value of information sets. Value can be very difficult to estimate. Some organizations have made such estimates, but most (especially small and medium sized organizations) have not. A conservative use of the model requires that the organization proposing to use the model provide some confidence level and margin of error attached to the value estimates. Additionally, the model's developers need to quantify how much the confidence level and margin of error of the output vary as a function of the confidence level and margin of error of the inputs.

Vulnerability of information sets. The vulnerability can also be very difficult to estimate. Because there are no reliable methods for estimating vulnerability, its value depends, among other things, on the changing threat landscape, on whether the particular organization is a favorite target, and on the architecture and access protocols of the information systems being protected. Transparency requires that the methods used to

make these estimates be spelled out, by the users if not by the model makers, and a level of confidence and margin or error must be assigned to them.

Form of the security function. Gordon and Loeb make a number of assumptions about the security function. Then, they prove that for two classes of functions meeting their assumptions, the optimal amount to invest to protect an information set is no more than 1/*e* (roughly 37%) of the potential loss from a successful attack. They conjecture that the 1/*e* fraction applies to all security functions meeting their assumptions. Willemsen (2007) describes a function meeting Gordon and Loeb's assumptions that forces expenditures of close to 50 percent of the potential loss. Further, by relaxing the assumptions slightly in a natural way, Willemsen shows that there are security functions that result in optimal spending levels close to 100 percent of the potential loss. This demonstration illustrates the potentially dramatic effects of simplifying mathematical assumptions.

To enable extension of their model from the two specific classes of functions to more general use, two essential questions must be answered:

- What reason is there to believe, intuitively or empirically, that these function classes capture a significant fraction of real-world situations?
- In what contexts have the classes been used before? Are they common to economic analyses? Have they been used to good effect in the past?

Without good empirically-based answers, the model cannot be extended with any confidence in the results.

The outputs from the accounting model are:

- The marginal rate of return on additional security investment to protect any given information set. Return is defined as increased security.
- The optimal amount to invest in securing a given information set, defined as the (unique) point where the marginal rate of return drops to zero.

Discussion of outputs:

Marginal rate of return. Needless to say, the accuracy of the marginal rate of return output by the model depends on the accuracy of the inputs and the fidelity of the security function. We discussed the assumptions about the mathematical form of the function above, when we addressed the model's simplifying assumptions. However, there are additional questions to be answered, about what data are relevant to the formulation of the security function. What affects the rate of return? The model assumes that the primary factor is threat reduction. Approximating this reduction is as difficult as estimating the baseline level of threat, or perhaps more so. But even after the reduction is estimated, there are additional factors to consider. For example, suppose that, based on the an initial formulation of the security function, a decision-maker decides to spend nothing. In this case, the model assumes that the threat level will not change. In some cases, this assumption may be true. However, there are many scenarios in which such an assumption may be badly flawed. For instance, when an organization or its product is highly visible, there may be a great deal of public scrutiny of the resources it devotes to security. In such a situation, if the organization spends nothing, it may acquire a

reputation for having bad or inadequate security. As a result, attacks may increase, because the organization is a more appealing target than those that are perceived as actively addressing their threats. In addition, the organization may lose market share as customers become concerned about security. These secondary effects generated by zero or low spending levels must somehow be taken into account when the model is used. There is more than one way to accomplish this. The model can be revised to force levels of spending to exceed a certain threshold. Or the model can be used in conjunction with constraints derived from considerations external to the model. In any case, fidelity of the model to the real world requires extreme care in formulating the security function and its constraints. For transparent application of the model, the factors that go into formulating the security function must be made explicit.

Optimal spending level. One of the key assumptions about the security function is that it is continuous, increasing, and differentiable. This description implies that any incremental increase in spending yields an incremental increase in security. There are clear counterexamples to this, the most simple being the organization that increases spending with an increasing threat, just to maintain a given security level. That is, the organization is "treading water," spending money just to stay in the same security place. For this reason, it is necessary to explain under what conditions a security function meeting Gordon and Loeb's assumptions is a reasonable approximation of the real world. For example, suppose that, because of the nature of the threats, security spending must be made in large, discrete chunks to be effective. Such a situation may obtain if a security patch must be applied to every workstation in a very large organization. In this case, a small spending increase may yield nothing at all in the way of increased security, whereas a large spending increase may yield substantially better security. Here, the security function is more akin to a step function than to a smooth curve, and the mathematics of the Gordon and Loeb model do not work out. In particular, there may be more than one point at which the marginal rate of return is zero; the model's claim of a single optimum spending level that can be derived from the marginal rate of return is no longer true. The model's users need to understand how sensitive the model's output is to the form of the security function. More generally, for any model, when the function fails to adequately reflect the world situation, is the output always misleading? Can it sometimes be informative despite the lack of the security function's fidelity?

Assumptions about the security function

Based on our discussion of the security function, we can make explicit several of its assumptions:

- Security, or expected security, is an increasing function of resources expended.
- In particular, if no resources are expended, security will not worsen.
- The security function is continuous and differentiable.
- The probability of security failure asymptotically approaches zero as investment increases.

Other assumptions

There are other, related assumptions that should also be made explicit:

- The resources available to any individual or organization to spend on information security have a finite upper bound. Similarly, the potential loss due to failure of information systems is limited.
- A decision about security investment can be made by looking at the security environment at a single moment in time. In other words, the model is a one-period model.
- There is a single threat to each information set.
- Security investments can reduce vulnerability but not the degree of threat.

Discussion of assumptions

We discuss these assumptions in turn, because changes to each one can have dramatic effects on the model's output.

The security function is continuous, increasing, and differentiable. One implication of these assumptions is that any money, time, or other resources devoted to improving security will in fact result in a net improvement in security. We discussed above one problem with assuming that incremental spending increases yield incremental security increases. But the basic assumption of an increasing security function is open to question. It is sometimes the case that money spent on security actually decreases security. For instance, if, for whatever reason, the security mechanisms purchased or implemented do not work as intended, the resulting security may actually worsen. Moreover, as with any attempt to improve software quality, the probability of system failure can actually increase when a security fault is fixed. For example, patching a security hole can uncover another fault that was not previously enabled or could not be executed. Such situations are surprisingly common, as is demonstrated when short inter-failure times occur, even at the end of data sets where long-term reliability is increasing. (Musa, Iannino and Okumoto 1990)

Thus, before a model of this type can be applied, two steps should be taken:

- The organization should make some estimate of whether its resource expenditures have in the past improved security consistently.
- If security does not turn out to be reliably increasing with expenditures, the model should be checked to see whether the conclusions drawn from its application fail if the security function is not monotonically increasing.

If the successful application of a model is highly sensitive to the shape of the security function, further analysis is needed. It may be that another type of function can be substituted, and a modified version of the model still used successfully. Or it may be that the model simply will not work, and another model should be sought.

If no resources are expended, security remains constant. This assumption is a way of isolating the analysis of the part of security that an organization can control from the part that it cannot control. Security can worsen for many different reasons—new, more dangerous threats emerge, a company becomes a favorite target of expert hackers, physical infrastructure is damaged by a natural disaster, and so on. This change may be no problem when applying a model; the baseline level of security is simply shifted downward. In some cases, however, a decrease in security may be indirectly related to

the spending level in a way that is not captured in the security function. As discussed earlier, spending levels can affect a company's reputation, which in turn can affect its security. Organizations should have some sense of whether such considerations are relevant for their situation. If so, the security function should be altered to capture them, if the model so allows. If not, the model must be supplemented by constraints arising from the interaction among spending levels, reputation, and security, or other constraints external to the model.

Given unlimited resources, the probability of security failure can be made arbitrarily close to zero. This is an example of a simplifying assumption that is almost always false. But the falsity may not actually invalidate the model's output. Why is it generally false? Because in the real world, given existing security technology, the measures needed to reduce the probability of security failure to near-zero levels would in almost every case make it impossible for an organization to perform its core functions. Nonetheless, the optimum spending level output by the model may not be sensitive to the function's behavior at much higher spending levels, so there is no harm in making this assumption.

Both resources and potential damage or loss have finite upper bounds. These assumptions limit application of the model to non-global, non-catastrophic situations. In the case of a national emergency, both spending levels and potential losses are effectively unlimited.

Security investment decisions can be made on the basis of a one-time snapshot of the security situation. It is perhaps inappropriate to call this an assumption of the model. The model outputs an optimal spending level based on a one-time evaluation of the security situation. It does not actually assume that the output can be used as the basis for more than a short-term spending plan. Using the model effectively, in any case, requires running it iteratively at regular intervals, as often as is required to adjust to dynamic changes in an organization's security environment. Depending on the volatility of the security environment, it may be impossible to base long-term spending plans on this model, even if it is used iteratively.

There is a single threat to each information set. Most organizations concerned about security have a threat model that makes clear what are the likely threats, how serious they are, and what might be done in response to each. To understand how the nature of the threats affects the output of the accounting model, consider two examples. First, suppose some capabilities are outsourced in order to shift risk away from the organization. That shift changes both the security function and its inputs. Second, suppose that there are multiple threats, each requiring different types of security measures. This heterogeneity may force security spending to be made in large discrete chunks. As we have noted, in this situation, the security function may be discontinuous.

Security investments can reduce vulnerability but not threat. There are ways that actions by an organization may change the level and type of threat. Investing in reputation may change the desirability of attacking the organization (up or down). Becoming more secure may discourage recreational attackers but encourage those attackers who enjoy an increased challenge. Some of these preventive investments are distinct from other security spending (e.g., reputation development); others coincide (e.g., when becoming

more secure changes the threat model). Using the model accurately requires somehow taking into account changes in threat consequent on security spending.

Decisions the accounting model supports:

- What fraction of the potential loss caused by a successful attack on an information set should be spent to secure it?

Discussion: The fraction of potential loss to invest is determined by the shape of the security function alone, so that determination may be made without good estimates of the value and vulnerability of information assets. If the model is to be used to determine actual spending levels, then the accuracy of available estimates of value and vulnerability comes into play.

Applicable domain of the accounting model:

- An organization that has information of value and resources to protect it, and in addition is able to determine
    - the value of its information,
    - the vulnerability of its information, and
    - that the security function relating its spending to its security level can be approximated by a function meeting the model's assumptions.

It is difficult to find empirical data as evidence that the accounting model accurately represents the phenomena of interest. That is, the model has many simplifying assumptions and is embedded in the larger context of how each security fix affects the overall software's quality. This dependence on the security function must be tested using a variety of sensitivity analyses, to determine how the assumption of increasing security affects the resulting recommended security investment. Moreover, there is no evidence that the model's applicability in the two examined cases allows interpolation to other situations.

Summary of Accounting Model Characteristics

The accounting model is intended to give guidance about how much an organization should spend to protect the value of a given set of information. Its results are meant to be general in form. The model user can explore what bounds can be placed on security investment, given that the security function is a member of some class specified in the model. Its usefulness depends critically on the fidelity of the security function, the accuracy of input quantities, and information sets being defined in such a way that the investments made to support each set are not interdependent. It cannot be used in practice unless there are data available allowing organizations to determine the shape of their security functions (functions relating investment to increased security), to calculate the value and the vulnerability of their information, and to classify information so that it may be divided into sets appropriate for use of the model.

Game-theoretic Model

The inputs to the model are:

- Demarcation of systems and subsets of players supporting them

- Cost per unit of effort for each player (effort means any resource expenditure)
- Value (benefit) to each player of the system operating successfully
- Mathematical form of the security function
- Mathematical form of the input to the security function
- Quantities to be optimized

Discussion of inputs:

Demarcation of systems and players. Just as the Gordon and Loeb accounting model placed constraints on its information, requiring division of information into discrete subsets, this model requires defining a single system and its supporting resources. If there is only one system in the universe in question, this definition is straightforward. If, on the other hand, there are overlapping or interconnected systems whose users and supporters also intersect, then we again have several questions to ask: What criteria can and should be used to demarcate systems and the agents (players) expending resources to support them? Which criteria preserve the validity of the model output? Which render it invalid? How (and how easily) may the game be restructured to accommodate different criteria for demarcating systems?

Costs and benefits. Both cost and benefit are quantities that may be difficult to estimate. If the cost is expressed in terms of money spent, it may be easy to determine. But if it includes time devoted by employees, opportunity cost, and other resources, calculation is harder. The value of the system's successful operation is a quantity similar to the value of the information set in the accounting model. If the model were intended to be used to calculate optimal individual expenditures, the same issues regarding confidence and accuracy would apply here as in the accounting model. However, the model is meant to be used to explore differences in optimal distribution of expenditures that result from optimizing different functions, representing different interests. For that purpose, actual costs and benefits are not required. In fact, assumptions about costs and benefits are treated as parameters. The model explores what happens as assumptions about costs and benefits are varied. The particular cases Varian examines are these:

- Case 1: All costs are identical
- Case 2: All costs and values are identical
- Case 3: Costs and values are drawn from a probability distribution (so that the maximum and minimum cost and value become more extreme as the number of players increases, while the expected cost and value remain constant).

Similarly, the model is used to explore various policies (such as imposing fines) that may be used to close the gap between private and public interests. It does not determine the dollar amount of fines, just the situations in which different penalty structures will be effective.

Form of the security function and its input. In the accounting model, the security model takes a dollar amount as input, and outputs an increase in security. In Varian's game-theoretic model, the security function takes as input the number of units of effort expended. It outputs the probability that the system will operate successfully. Describing

the relevant probability function is not easy; there is no reliable method to estimate the probability that a system will fail. As with the accounting model, some level of confidence and margin of error must be assigned to the function outputs, and the sensitivity of model outputs to those quantities must be carefully characterized.

Unlike the security function in the accounting model, the security function in the game-theoretic model does not take as input resource expenditures input by the user. Here, the resource expenditures of individual players are *output* as equilibrium strategies. What the user contributes is a function expressing how the security of the system depends on the contributions of individual players. In this particular model, this expression can be a version of one of three prototypical forms, or some combination thereof. These are:

- Total Effort: The security function takes the sum of all effort expended as input.
- Weakest Link: The security function takes the minimal effort expended as input.
- Best Shot: The security function takes the maximal effort expended as input.

The user of the model must determine what the correct function is. This determination may or may not be easy; in any case, guidance is needed about what kinds of simplifications or approximations of a real-world function result in approximately correct outputs, and which render the model invalid. At the other extreme, guidance is needed about how complex the function can be made without rendering the model computationally intractable.

Quantity to be optimized. The games used in this model have equilibrium points for the set of individual strategies, where each agent is trying to optimize its own net benefit. However, the model is not intended to be used by individual players to calculate how much effort to expend. Rather, it is intended to be used to compare the equilibrium strategies with the distribution of expenditures that would result if the *social good* (defined as the total net benefit, summed over all players) were optimized. That sum is one function a policy maker might wish to optimize. The model can equally well be used to explore the distribution of expenditures that result from optimizing other sets of interests.

The outputs from the game-theoretic model are:

- Equilibrium strategies giving optimal expenditures for each player
- Comparisons between equilibrium strategies produced by optimizing different functions
- Policies that will make the equilibrium strategies for two different optimization functions equal

Discussion of outputs:

Comparison of different optimization functions. As noted above, functions can be chosen to represent any combinations of interests. The resulting expenditure distributions can be compared. Suppose a policy-maker wishes to optimize over a function *F*, and there is a gap between *F* and the Nash equilibrium (where each player optimizes its own net benefit). The model may be used to explore policies that impose new costs and/or benefits, in order to eliminate the gap. The output is *not* the actual equilibrium strategies

with dollar amounts attached; rather, the output is a comparison among optimization functions.

Policies. The model enables the user to explore various policy types derived from standard legal models (fines, liability, negligence, due diligence). Some of the policies that work mathematically may be entirely unworkable from a practical point of view. For example, one policy considered in Varian's paper makes the player with the lowest marginal cost of effort pay a fine if the system fails, while the others pay nothing. This policy might well violate some players' sense of fair play to the point that they withdraw from the game. One of the authors of this paper once heard a talk, delivered in all seriousness by a professional ethicist, arguing that the best and most cost-effective way to reduce crime would be to sell permits allowing people to commit individual crimes, priced according to the severity of the crime. It need hardly be said that the audience was unreceptive, and not because of flaws internal to the game-theoretic argument.

The assumptions made by the game-theoretic model are:

Assumptions about the security function

- Security, or expected security, is an increasing function of resources expended.
- In particular, if no resources are expended, security will not worsen.
- The security function is continuous and differentiable.

Other assumptions

- The resources available to any individual or organization to spend on information security have a finite upper bound.
- A decision about security investment can be made by looking at the security environment at a single moment in time. (The game is a single round, whether simultaneous or sequential.)
- Every additional dollar (or unit of effort) spent yields the same amount of additional security.
- System failure is Boolean: Either the system is working, or it has failed. That is, the model does not allow for degrees of failure.
- All individual agents have perfect information. In particular, each individual knows the costs and payoffs of all agents involved. In addition, each agent relies on the same security function in calculating optimal investment.
- Public or Social Good is optimized when the total net benefit, summed over all agents, is maximal.

Discussion of assumptions:

Finite upper bound on resources. This assumption effectively limits the model to scenarios where no player has resources incomparably greater than the rest.

Security investments made on the basis of a one-time snapshot of the security environment. This assumption is common to the game-theoretic model and the Gordon and Loeb model. The same comments apply.

Every dollar buys the same amount of security. By contrast, the Gordon and Loeb accounting model assumes that the marginal value of additional investment decreases as the total investment increases: The more money you spend, the less you get for it. The difference between these two assumptions highlights the importance of getting the security function right. Gordon and Loeb's results would be meaningless if the marginal value of additional security investment were constant, as is assumed here.

System failure is Boolean. This assumption limits the applicability of the model to scenarios where failure does not happen in degrees: Either the system works, or it crashes. In most real-world situations, the resources expended to prevent total failure also serve to mitigate the damage. For the insights derived from the model to be applicable, the costs and benefits associated with both lower levels of failure and total system failure must be taken into account.

The game is a perfect information game. This assumption is rarely true in the real world; depending on how reasonable an approximation the perfect information assumption is to the real-world situation under consideration, the model may have to be extended to a game without perfect information in order to be applicable.

The public good is equal to total net benefit. The public good can be measured in a variety of ways. Here it is assumed to be measured by the sum of the payoffs for all players, less the sum of all costs. There may be other definitions preferred by decision-makers in some situations. For example, it may be desirable to maximize the minimum net benefit, to minimize the difference between the minimum and maximum net benefits, or to maximize the average net benefit. A decision-maker applying the model should make a conscious choice about which function to use to represent the public good. In most cases, the model can accommodate an alternate function, but it should be checked to make sure.

Decisions the game-theoretic model supports:

- How to structure policies imposing cost penalties and incentives to maximize a particular set of interests.

Discussion: The model can be used to show how optimizing different sets of interests results in different distributions of investment levels in support of a system. It can also test strategies for altering costs and benefits to change those outcomes. Policy-makers can use the model (provided its assumptions are met) to determine how to structure penalties and incentives. Determining actual amounts of penalties and incentives would require accurate estimates of existing costs and benefits.

Applicable domain of the game-theoretic model:

The model may be used to explore penalty and incentive strategies to further the interests of any subset of a group of agents all of whom benefit from the successful operation of a system, and contribute varying amounts to support it.

As with the accounting model, it is difficult to find empirical data as evidence that the game-theoretic model accurately represents the phenomena of interest. We have seen how many of the model's assumptions (such as perfect information) may be unrealistic, and we do not know the game-theoretic model's sensitivity to attempts to relax them. In

addition, the dependence on the security function must be tested using a variety of sensitivity analyses. Moreover, it is not clear whether every situation can be described as a single system, as required by the model.

Summary of Game-theoretic Model Characteristics

The game-theoretic model is intended to explore situations in which a system is supported and used by a number of individual players whose interests do not necessarily coincide. Given assumptions about the shape of the security function relating security spending to increased security (or reduced vulnerability), the game-theoretic model may be used to understand what kinds of situations result in an undesirable distribution of expenditures and benefits. The model provides a way to capture what a desirable distribution would look like. As with the accounting model, the usefulness of the game-theoretic model depends critically on the shape of the security function. If data are not available to determine what the real security function looks like, the model's insights are not applicable. An additional requirement is philosophical rather than empirical. Assigning desirability to a distribution of expenditures and benefits is an expression of whose interests, or what group's interests are most valued, and how they are best served. The desirability is not always a matter of consensus, and it is non-trivial to formulate. The law of unintended consequences operates forcefully in this domain.

Input/Output Model

Our discussion of the input/output model is necessarily much more general than our discussion of the accounting and game-theoretic models. The input/output model is presented speculatively as an illustration of how research into the effect of intellectual property theft on the U.S. economy might profitably be conducted. A great deal of economic data is available to feed an input/output model of the U.S. economy, but much of the data required for this specific use of input/output modeling does not yet exist. In addition, since the model is macro-economic and global, there are many extra-economic factors (e.g., politics and culture) affecting the domain under investigation.

The underlying model used by Andrijcic and Horowitz is a Leontieff linear equilibrium model of the U.S. economy. (Leontieff, 1951) Santos and Haimes have adapted a general input/output model to investigate the "inoperability" caused by reduction in demand due to terrorism. (Santos and Haimes, 2004) "Inoperability" is defined as the "level of a system's dysfunction expressed as a percentage of its 'as-planned' production capacity."

The Leontieff input/output model represents the interdependence of all sectors of the U.S. economy, expressed in terms of each sector's demand for the output of other sectors, and production output that is supplied to other sectors. It is an equilibrium model, meaning that it represents the static case where supply is equal to demand. The model makes use of data from the U.S. Bureau of Economic Analysis, which compiles and periodically publishes data describing the relationships among the sectors.

The model is linear. Santos and Haimes posit a perturbation in demand for the products of one or more sectors, and then solve to determine the resulting economic loss. They restrict themselves to a scenario where the direct result of terrorism is a reduction in demand. This approach assumes that terrorism's effect on the economy is primarily due to the fear it causes, leading to reduced demand. For example, after the terrorist attacks

on September 11, 2001, demand dropped dramatically in the commercial aviation sector. They do not model effects on production caused by direct destruction of infrastructure or production facilities.

In the inoperability model, the model may be used to investigate the interactions among sectors of the U.S. economy for a range of input parameters. The model suffers from an assumption underlying all linear models: that changes in inputs and outputs are small and bounded with respect to the absolute values of the input parameters. In general, ten percent is the maximum perturbation that can be considered valid in a linear system context, absent other information regarding the dynamics of the system. Santos and Haimes investigate what happens if there is a ten percent reduction in demand for the air transportation sector.

Andrijcic and Horowitz combine the Santos-Haimes interoperability model with two others. The domain they are investigating is the interdependence of U.S. firms and their international supply chains. Many of these supply chains are in countries that are commonly the source of intellectual property theft. The premise of the model is that due to probability of theft of intellectual property, there is increased risk that U.S. firms will lose market share to international competitors, and that these losses will be long-lasting.

The first component added to the input/output model is a model of the equity of major private sector firms in a particular sector. The authors assume that the loss of intellectual property will result in a loss of equity in U.S. based firms. The equity model calculates the average market capitalization of a firm in a particular sector and uses that as a proxy for average loss due to the theft of intellectual property. The argument is that a publicly released theft of intellectual property will reduce the long term earning prospects of the firm and lead to a loss of equity value.

The second component is a foreign market share multiplier, quantifying how much of the world market in a particular sector already resides in a foreign country. Andrijcic and Horowitz. caution that both of these additional models are preliminary and require additional analysis and validation.

An additional component of the model is an Espionage Propensity Factor (EPF), quantifying the likelihood that a particular country will engage in corporate espionage. The EPF is dependent on a number of parameters, including foreign market share, the imports and exports in a sector, and the supply base for U.S. firms in a particular sector.

Using these components, the input/output model is used to identify sectors that are vulnerable to intellectual property theft. For these sectors, Andrijcic and Horowitz calculate the loss to the U.S. in terms of the equity that U.S. firms lose to foreign firms.

All macro-economic models involve enormous simplifications in order to arrive at usable representations. They are consequently poor at making predictions. Since this model involves assumptions about the propensity of firms in other countries to commit certain kinds of crimes, as well as assumptions about the behavior of the public in light of information about theft, there are many relevant factors that either cannot be measured or are exceedingly difficult to measure (or even to perceive). The usefulness of such a model is primarily to identify, in general terms, which sectors are most likely experience

losses due to intellectual property theft, assuming that international patterns of theft remain more or less constant.

Conclusion

As industry and government seek ways to balance investment in cyber security with other demands on resources, decision support tools and techniques are needed to frame the problem and convey important information and relationships. Many of the tools and techniques rely on underlying models. Since any model is suitable for some purposes but not others, we have presented principles and a framework with which each model can be evaluated. Using accounting, game-theoretic and input/output models as examples, we have shown how the principles and framework can be employed. Such an evaluation can not only suggest the appropriateness of each model but also highlight where new or modified models are needed for addressing gaps or additional uncertainties. This work is preliminary; we plan to extend the use of the principles and framework to a large set of cyber security economics models, to help depict the landscape of models available to decision-makers. At the same time, we are investigating the utility of existing datasets that inform the models. Because a model is useless without good data, it may be necessary to simplify and tailor the models to the available data, rather than build more elaborate models for which data are neither representative nor credible.

References

Andrijcic, Eva, and Horowitz, Barry. 2004. A Macro-Economic Framework for Evaluation of Cyber Security Risks Related to Protection of Intellectual Property. Workshop on the Economics of Information Security, Cambridge, England.

Baer, Walter S. 2003. Rewarding IT security in the marketplace. Santa Monica, California: RAND Corporation.

Baer, Walter S. and Andrew Parkinson. 2007. "The Role of CyberInsurance in Managing IT Security" *IEEE Security and Privacy Magazine* (May/June).

Campbell, Katherine, Lawrence A. Gordon, Martin P. Loeb, and Lei Zhou. 2003. The economic cost of publicly announced information security breaches: empirical evidence from the stock market. *Journal of Computer Security* 11:431-448.

Conrad, James R. 2005. Analyzing the risks of information security investments with monte-carlo simulations. Paper read at IEEE Workshop on the Economics of Information Security.

Dynes, Scott, Hans Brechbuhl, and M. Eric Johnson. 2005. Information security in the extended enterprise: some initial results from a field study of an industrial firm. Hanover, New Hampshire: Tuck School of Business, Dartmouth College.

Farahmand, Fariborz, Shamkant B. Navathe, Gunter P. Sharp, and Philip H. Enslow. 2005. A Management Perspective on Risk of Security Threats to Information Systems. *Information Technology and Management* 6 (2-3):203-225.

Gal-Or, Esther, and Anindya Ghose. 2005. The economic incentives for sharing security information. *Information Systems Research* 16 (2):186-208.

Garcia, Alfredo and Barry Horowitz. 2006. The Potential for Underinvestment in Internet Security: Implications for Regulatory Policy. Workship on the Economics of Information Security, Cambridge, England.

Geer, Daniel E. 2001. Return on security investment: calculating the security investment equation. *Secure Business Quarterly* 1 (2).

———. 2004. Security of information when performance matters. Waltham, Massachusetts: Verdasys, Inc.

The GHG Protocol for Project Accounting. 2005. Washington, District of Columbia: World Business Council for Sustainable Development; World Resources Institute.

Gordon, Lawrence A., and Martin P. Loeb. 2002. The Economics of Information Security Investement. *ACM Transations on Information and System Secuity* 5 (4):438-457.

Gordon, Lawrence A., and Martin P. Loeb. 2005. Economic aspects of information security. College Park, Maryland: The Robert H. Smith School of Business, University of Maryland.

Gordon, Lawrence A., Martin P. Loeb, and Tashfeen Sohail. 2003. A framework for using insurance for cyber-risk management. *Communications of the ACM* 46 (3):81-85.

Haimes, Yacov Y., and Clyde G. Chittester. 2005. A roadmap for quantifying the efficacy of risk management of information security and interdependent SCADA systems. *Journal of Homeland Security and Emergency Management* 2 (2).

Horowitz, Barry, and Alfredo Garcia. 2005. A growing trend towards underinvestment in internet security. Charlottesville, Virginia: University of Virginia, Department of Systems and Information Engineering.

Hull, John C. 1997. *Options, Futures, and Other Derivatives*. Third ed. Upper Saddle River, New Jersey: Prentice-Hall.

Irvine, Cynthia E., and Michael F. Thompson. Expressing an information security policy within a security simulation game. Monterey, California: Naval Postgraduate School.

Irvine, Cynthia E., Michael F. Thompson, and Ken Allen. 2005. CyberCIEGE: Gaming for information assurance. *IEEE Security and Privacy Magazine* (May/June):61-64.

Johnson, M. Eric and Eric Goetz. 2007. "Embedding Information Security Into the Organization. *IEEE Security and Privacy Magazine* (May/June).

Leontieff, W.W. 1951. The Structure of the American Economy, 1919-1939: An Empirical Application of Equilibrium Analysis, 2nd Edition. New York, NY: Oxford University Press.

Morgan, M. Granger, and Max Henrion. 1990. *Uncertainty: a guide to dealing with uncertainty in quantitative risk and policy analysis*. Cambridge, United Kingdom: Cambridge University Press.

Musa, John D., Anthony Iannino and Kazuhiro Okumoto. 1990. *Software Reliability: Measurement, Prediction, Application*. New York: McGraw-Hill.

Pfleeger, Shari Lawrence, Martin Libicki and Michael Webber. 2007. "I'll Buy That! Suggestive Findings About Cyber Security in the Internet Marketplace" *IEEE Security and Privacy Magazine* (May/June)

Santos, Joost R., and Yacov Y. Haimes. 2004. Modeling the demand reduction input-output (I-O) inoperability due to terrorism of interconnected infrastructures. Charlottesville, Virginia: University of Virginia, Department of Systems and Information Engineering.

Soo Hoo, Kevin J. 2000. How much is enough? A risk-management approach to computer security. Palo Alto, California: Stanford University, Consortium for Research on Information Security and Policy.

Varian, Hal R. 2004. System Reliability and Free Riding. University of California, Berkeley.

Willemson, Jan. 2006. On the Gordon and Loeb Model for Information Security Investment. Workshop on the Economics of Information Security, Cambridge, England.