

Optimal Allocation of Resources for Defense of Simple Series and Parallel Systems from Determined Adversaries

Vicki M. Bier¹ and Vinod Abhichandani²

Abstract

Managing the risks posed by an intelligent and adaptable adversary is different from many other types of risk management. Thus, risk management in this context is a problem of game theory rather than decision theory. In other words, the defender wishes to choose the optimal strategy for defending against an optimal attack, and vice versa. In this paper, we apply game theory to help in characterizing optimal defensive strategies against intentional attacks. The results yield useful insights.

Introduction

In the aftermath of the September 11th, 2001, attacks on the World Trade Center and the Pentagon (and the anthrax attacks in the United States), there is increased interest in strategies for protecting assets of value (including human life) against attacks by an intelligent and adaptable adversary. Even before that time, there were calls for greater attention to critical infrastructure protection, including computer security; see, e.g., President's Commission on Critical Infrastructure Protection [1997].

This is a fundamentally different challenge from protecting against “acts of nature” or “accidents.” For example, an earthquake will not become stronger or “smarter” just because we have hardened our buildings to protect against it. In contrast, an intelligent and determined adversary is likely to adopt a different offensive strategy once we have put a particular set of protective measures in place. Therefore, good defensive strategies must consider the adversary's behavior.

¹Professor, Dept. of Industrial Engineering, University of Wisconsin-Madison, 1513 University Avenue, Madison WI 53706; 608-262-2064; bier@engr.wisc.edu

²Graduate student, Dept. of Industrial Engineering, University of Wisconsin-Madison, 1513 University Avenue, Madison, WI 53706; 608-263-2687; vinod_abhi@yahoo.com

To illustrate, if we harden one point of entry into a system to make it essentially invulnerable, an adversary is likely simply to target an alternative point of entry instead. Thus, Schneier [2001] notes: “The defender has to defend against every possible attack. The attacker, on the other hand, only has to choose one attack, and he can concentrate his forces on that one attack.”

The requirement to take into account the behavior of adversaries means that we must consider their goals and motivations when we select defensive strategies [Schudel and Wood 2000]. One of the most important characteristics may be whether an adversary is opportunistic or determined. Opportunistic adversaries may simply be looking for easy targets. In other words, they may be interested in a wide range of targets and shift to an easier one if their current target proves too difficult or costly to attack. Ordinary vandalism typically falls into this category; similarly, many computer hackers may not care which internet businesses they disrupt. In contrast, realistic levels of difficulty or cost will not necessarily deter a determined attacker. Examples of determined attackers might include military adversaries or terrorists, unethical business competitors, or even disgruntled ex-employees who are determined to damage a company’s assets or reputation.

Defending against opportunistic adversaries is less difficult than defending against determined attackers. As in the old joke about the two campers who see a bear, in defending against opportunistic attackers you do not have to run faster than the bear; you just have to run faster than the other potential victim does! Thus, if an organization is convinced that most of its adversaries are opportunistic, it may be acceptable merely to adopt security practices that are better than the average for similar companies or systems. In this type of situation, the organization’s defenses need only be strong enough to discourage possible attacks.

Defending something of value from a truly determined attacker is much more difficult. To illustrate some of the challenges involved, we will compare the optimal strategies for defending simple systems (configured either in series or in parallel) against attacks by knowledgeable and determined adversaries.

General Problem Formulation and Notation

Consider a system made up of multiple components. Redundant components that provide alternative means to perform the same task are functionally parallel to each other. Similarly, components that are all necessary to perform a single task are functionally in series. Of course, real-world systems may consist of a large number of components in complex combinations of parallel and series configurations. As a building block toward understanding optimal strategies for investing in the security of such systems, we discuss simple systems with n components, either in parallel or in series with each other, subjected to limited attacks by an intelligent adversary.

To derive a better understanding of the security of such systems, we first develop mathematical formulations for systems with only two components, corresponding to various possible assumptions regarding attacker and defender knowledge, constraints, and objectives. We then extend the results to systems with more than two components. We assume that the attacker wishes to maximize the probability of success for an attack on the system. In addition, we assume that

attacks against different components succeed or fail independently of each other. The assumption of independence is obviously somewhat restrictive; it is likely to be more nearly satisfied in systems that have already invested in basic reliability and security measures such as spatial separation and functional diversity.

We summarize below some mathematical notations and symbols used throughout the paper:

C : The total budget constraint (where applicable).

C_i : The investment allocated to defend component i .

$P(C_1, C_2, \dots, C_n)$: The probability of success of an attack against the system, as a function of the defensive investments in all components.

$P_i(C_i)$: The probability of success of an attack against component i , as a function of the resources C_i expended to strengthen that component. The functions $P_i(C_i)$ are deemed to be convex, decreasing, strictly positive, continuous, twice differentiable, and invertible (with differentiable inverses).

α : The expected dollar value of reducing the probability of success of an attack against the network from one to zero; e.g., the dollar value of the system if operational, multiplied by the probability of an attack against the system.

$F(C_1, C_2, \dots, C_n)$: The objective function that the defender wishes to minimize in any given case. In the case where the defender has a budget constraint, this will simply be the probability of success of an optimal attack against the system. In the case with no budget constraint, the objective function will be a weighted sum of the probability of success of an optimal attack against the system and the defensive investments made by the defender, with the weight given by the parameter α .

Components in Parallel

To begin, consider a system of two components in parallel. Since the components are redundant, either one is sufficient to ensure successful operation of the system. Therefore, attackers with the goal of disabling the network would have no choice but to attack and disable both components, even though they may have a choice about which component to target first. We assume that the defender is concerned about the probability that an attack will succeed, but not the time at which the attack occurs, so it does not matter whether the two attacks occur simultaneously or sequentially.

If the defender spends C_1 on protection of component 1 and C_2 on protection of component 2, then under the assumption of independence, the resulting probability of an attack on the network succeeding will be given by $P(C_1, C_2) = P_1(C_1) P_2(C_2)$. We now develop mathematical models for optimal defense of this system for two possible cases, depending on whether the defender has a cost constraint.

Constrained case. Here, we assume that the defender has limited resources available for enhancing the system defenses, and wishes to maximize the system security subject to the given budget constraint. Therefore, the objective function to be minimized by the defender is simply the total probability of an attack on the system succeeding; that is,

$$F(C_1, C_2) = P(C_1, C_2) = P_1(C_1) P_2(C_2) \quad (1)$$

subject to the constraints $C_1 + C_2 \leq C$ and $C_i \geq 0$.

It is easy to show that the objective function is monotonically decreasing. Therefore, the constrained minimum must occur along the line $C_1 + C_2 = C$; thus, at the minimum, we will have $C_2 = C - C_1$. Substituting this into the objective function, differentiating with respect to C_1 , and equating the result to zero, we get:

$$\frac{\partial F(C_1, C - C_1)}{\partial C_1} = P_1'(C_1) P_2(C - C_1) - P_1(C_1) P_2'(C - C_1) = 0$$

Thus, if a local optimum exists for $0 < C_1 < C$, it must satisfy:

$$\frac{P_1'(C_1)}{P_1(C_1)} = \frac{P_2'(C - C_1)}{P_2(C - C_1)}$$

To determine whether a point satisfying this condition is a minimum, a maximum, or an inflection point, we take the second derivative of the objective function:

$$\frac{\partial^2 F(C_1, C - C_1)}{\partial C_1^2} = P_1''(C_1) P_2(C - C_1) + P_1(C_1) P_2''(C - C_1) - 2 P_1'(C_1) P_2'(C - C_1)$$

A point satisfying the first-order condition will be a local minimum if this second derivative is positive, a local maximum (along the constraint $C_1 + C_2 = C$) if the second derivative is negative, and inconclusive if the second derivative equals zero.

Note that the global minimum also can occur at a point where one of the C_i equals zero and the other is equal to C , even if this point does not satisfy the first-order condition given above. For example, one component may be so much more costly to defend than the other that the optimal defense of the system involves a zero expenditure on defense of the more costly component. An optimum with no investment in one component can occur under several conditions. First, if there is only one critical point along the constraint $C_1 + C_2 = C$, it may be a local maximum or an inflection point (when viewed as a one-dimensional function of C_1). Second, there may be more than one critical point along the constraint $C_1 + C_2 = C$. Finally, there may be no critical point along the constraint $C_1 + C_2 = C$.

Special cases.

1. Let $P_i(C_i) = a_i e^{-b C_i}$. Here, the first-order condition for optimality given above will be satisfied for all values of C_1 along the line $C_1 + C_2 = C$, so any point along this constraint will be a global minimum.
2. Let the $P_i(C_i)$ be log-convex (which implies, roughly speaking, that the probability of success of an attack against component i decreases faster than exponentially in the level of defensive resources invested). In this case, the

defender's objective function will be convex. Therefore, if a feasible point satisfying the first-order optimality condition exists, it will be the unique global minimum. However, if a feasible point satisfying the first-order optimality condition does not exist, the global minimum will occur where C_1 equals either zero or C , as before.

3. Let the $P_i(C_i)$ be identical, so that $P_1(C_1) = P_2(C_1)$ for all C_1 . Substituting this into the objective function from Equation (1), we get a new objective function, $F(C_1, C - C_1) = P_1(C_1) P_2(C - C_1) = P_1(C_1) P_1(C - C_1)$. To find the minimum, we differentiate with respect to C_1 and then equate the result to zero, giving $F'(C_1, C - C_1) = P_1'(C_1) P_1(C - C_1) - P_1(C_1) P_1'(C - C_1) = 0$, or:

$$\frac{P_1'(C - C_1)}{P_1(C - C_1)} = \frac{P_1'(C_1)}{P_1(C_1)}$$

As before, there can be multiple local optima, although now those optima must be symmetric. One critical point occurs at $C - C_1 = C_1$, or $C_1 = C/2$. However, this point need not be the global minimum, unless the objective function is convex (e.g., if the P_i are log-convex). For example, it may be more cost-effective to invest all defensive resources in a single component, in which case symmetric global minima would occur at $C_1 = 0$ and $C_1 = C$.

Extension to systems with more than two components. For a system with n components in parallel, the attacker must disable all components. Therefore, the probability of an attack succeeding against a system with n functionally redundant components is given by $P(C_1, C_2, \dots, C_n) = P_1(C_1) P_2(C_2) \dots P_n(C_n)$. To find the constrained minimum of this function, we again use the fact that the objective function is monotonically decreasing. Therefore, the constrained minimum must occur along the constraint $C_1 + C_2 + \dots + C_n = C$; i.e., at a point satisfying:

$$C_1 = C - (C_2 + \dots + C_n) \tag{2}$$

Substituting Equation (2) into the objective function yields a revised objective function given by $F(C_1, C_2, \dots, C_n) = P_1[C - (C_2 + \dots + C_n)] P_2(C_2) \dots P_n(C_n)$. To obtain the optimum of this function, we differentiate with respect to C_i for $i = 2, 3 \dots n$, and equate the results to zero, yielding:

$$\frac{P_1'[C - (C_2 + \dots + C_n)]}{P_1[C - (C_2 + \dots + C_n)]} = \frac{P_i'(C_i)}{P_i(C_i)}$$

and therefore in general:

$$\frac{P_i'(C_i)}{P_i(C_i)} = \frac{P_j'(C_j)}{P_j(C_j)} \tag{3}$$

Thus, if the global optimum occurs at a point where both C_i and C_j are non-zero, then at optimality, the percentage decrease in P_i caused by spending more on component i will equal the percentage increase in P_j caused by spending less on component j . As before, the optimum may involve spending no resources on one or more components (e.g., if they are much more costly to defend than others), in which case the first-order conditions given in Equation (3) may not be satisfied for pairs of components that do not both have positive defensive investments at optimality.

For example, consider a system made up of three components. For the reasons discussed above, the optimum must lie at $C_1 + C_2 + C_3 = C$. Figure 1 below graphically depicts the possible values of the C_i at optimality. If the global optimum occurs at an interior point of the triangle described by the points $C_1 = C$, $C_2 = C$, and $C_3 = C$, the first-order conditions given in Equation (3) will be satisfied for all pairs of the variables. If the global optimum occurs along an edge of the triangle (with one of the C_i zero), then only the two non-zero variables must satisfy Equation (3). Finally, the global optimum may occur at a vertex of the triangle (where one of the C_i equals C), in which case the first-order conditions in Equation (3) may not be satisfied for any of the three variables.

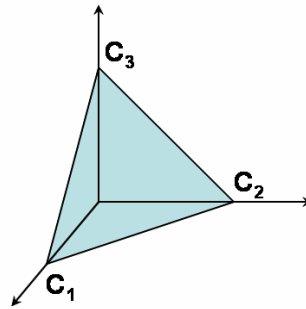


Figure 1. Possible variable values at optimality.

Interpretation of results. $P_i'(C_i)$ is the rate of change of the probability of success of an attack on component i . Dividing the rate of change by $P_i(C_i)$ essentially converts an absolute rate of change to a fractional rate of change. Therefore, the first-order conditions in Equation (3) essentially mean that if C_i and C_j are both positive at optimality, then the percentage decrease in P_i caused by spending more on component i will equal the percentage increase in P_j caused by spending less on component j .

Since the objective function for this problem is not necessarily convex, there may be multiple local minima. Thus, the analysis above does not tell us very much about the location of the global minimum. However, the key point to note is that the location of the optimum (at least for those components for which $C_i > 0$ at optimality) depends critically on the cost-effectiveness of defending the various components, as given by the derivatives of the P_i . Thus, defenders have the flexibility to allocate preventive resources to get the “most bang for the buck,” and in particular the option of not defending some components at all (e.g., if they are too costly to defend).

Unconstrained case. Here, the defender can spend as much as is justified to enhance the system security. In this case, we assume that the defender wishes to minimize the expected dollar impact of an attack (in terms of the lost value of the system) plus the cost of the protective investments; i.e., $F(C_1, C_2) = \alpha P_1(C_1) P_2(C_2) + C_1 + C_2$. To find the local optima for this function, we apply partial differentiation with respect to C_1 and C_2 , respectively, and equate the first derivatives to zero, yielding:

$$\frac{\partial F(C_1, C_2)}{\partial C_1} = \alpha P_1'(C_1) P_2(C_2) + 1 = 0$$

$$\frac{\partial F(C_1, C_2)}{\partial C_2} = \alpha P_1(C_1) P_2'(C_2) + 1 = 0$$

Thus, if the global optimum satisfies $C_1, C_2 > 0$, then at optimality we must have:

$$P_1'(C_1) P_2(C_2) = P_1(C_1) P_2'(C_2) = -1/\alpha \quad (4)$$

which also implies (as in the constrained case):

$$\frac{P_1'(C_1)}{P_1(C_1)} = \frac{P_2'(C_2)}{P_2(C_2)}$$

In Equation (4), $P_1'(C_1) P_2(C_2)$ is the marginal effect of an incremental unit of expenditure on component 1 in reducing the probability of success of an attack on the network. Thus, at any local optimum satisfying $C_1, C_2 > 0$, the marginal effect of an incremental unit of expenditure is inversely proportional to the value of the system.

To determine whether a critical point satisfying Equation (4) is a minimum, we apply second-order differentiation to the objective function, yielding:

$$\frac{\partial^2 F(C_1, C_2)}{\partial C_1^2} = \alpha P_1''(C_1) P_2(C_2)$$

$$\frac{\partial^2 F(C_1, C_2)}{\partial C_2^2} = \alpha P_1(C_1) P_2''(C_2)$$

$$\frac{\partial^2 F(C_1, C_2)}{\partial C_1 \partial C_2} = \alpha P_1'(C_1) P_2'(C_2)$$

Therefore, the Hessian matrix for this objective function is:

$$H = \begin{vmatrix} \alpha P_1''(C_1) P_2(C_2) & \alpha P_1'(C_1) P_2'(C_2) \\ \alpha P_1'(C_1) P_2'(C_2) & \alpha P_2''(C_2) P_1(C_1) \end{vmatrix}$$

The determinant is given by $J = \alpha^2 P_1''(C_1) P_2''(C_2) P_1(C_1) P_2(C_2) - \alpha^2 [P_1'(C_1) P_2'(C_2)]^2$. Substituting in the values of $P_1'(C_1)$ and $P_2'(C_2)$ from Equation (4), we get:

$$J = \alpha^2 P_1''(C_1) P_2''(C_2) P_1(C_1) P_2(C_2) - \alpha^2 \{1/[\alpha^2 P_1(C_1) P_2(C_2)]\}^2$$

$$= \alpha^2 P_1''(C_1) P_2''(C_2) P_1(C_1) P_2(C_2) - 1/[\alpha P_1(C_1) P_2(C_2)]^2$$

A critical point satisfying the first-order conditions will be a local minimum when J is positive; i.e., when $P_1''(C_1) P_2''(C_2) > 1/[\alpha^4 P_1(C_1)^3 P_2(C_2)^3]$, or in other words, when the parameter α is sufficiently large. When α is small enough, then no critical point will be a local minimum, and the global minimum will occur with C_1 and/or C_2 equal to zero. As before, this can also occur when there are multiple critical points.

Special cases.

1. Let $P_i(C_i) = a_i e^{-b C_i}$. Here, the first-order condition for optimality holds for all values of C_1 and C_2 satisfying $C_1 + C_2 = \ln[-1/(\alpha a_1 a_2 b)]/b$, and any point on this line segment will be a global minimum. Thus, as α grows, the optimal level of defensive investment will grow, which is intuitively reasonable.
2. Let $P_i(C_i)$ be log-convex. As in the constrained case, the objective function will now be convex, ensuring the existence of a unique local minimum (possibly with C_1 and/or C_2 equal to zero, depending on the value of α).
3. Assume now that the functions $P_i(C_i)$ are identical, so that the objective function can be written as $F(C_1, C_2) = \alpha P_1(C_1) P_1(C_2) + C_1 + C_2$. In this case, if the global optimum has both C_1 and C_2 positive, then by the results derived above, they must satisfy:

$$\frac{P_1'(C_1)}{P_1(C_1)} = \frac{P_1'(C_2)}{P_1(C_2)}$$

Thus, there will exist a critical point with $C_1 = C_2$. However, this will not necessarily be the global minimum unless the objective function is convex.

Extension to systems with more than two components. In this case, the success probability of an attack is $P(C_1, C_2, \dots, C_n) = P_1(C_1) P_2(C_2) \dots P_n(C_n)$; the objective function is $F(C_1, C_2, \dots, C_n) = \alpha [P_1(C_1) P_2(C_2) \dots P_n(C_n)] + C_1 + C_2 + \dots + C_n$. To find the optimum, we differentiate the objective function with respect to C_i and equate the result to zero, yielding:

$$\alpha P_i'(C_i) \prod_{j \neq i} P_j(C_j) + 1 = 0 \tag{5}$$

or in other words:

$$P_i'(C_i) = -1/[\alpha \prod_{j \neq i} P_j(C_j)] \tag{6}$$

As before, there can be multiple critical points, and/or a global minimum at which one or more of the C_i are equal to zero and hence need not satisfy Equations (5) and (6). However, if α is sufficiently large that both C_i and C_j are non-zero at the global minimum, then at optimality the percentage rates of change of the success probabilities of attacks against those components will be equal, and we will have:

$$\frac{P_i'(C_i)}{P_i(C_i)} = \frac{P_j'(C_j)}{P_j(C_j)} = -1/\left[\alpha \prod_{i=1}^n P_i(C_i)\right]$$

In other words, the percentage rates of change in the success probabilities of attacks against components i and j will be equal, and will be inversely proportional to the expected losses due to attacks on the system.

As the value of α decreases, Equation (5) will hold for fewer components at optimality, and the optimal solution will be to invest no resources at all in the defense of some components. When α is sufficiently small, the objective function will be monotonically increasing for all non-negative values of the C_i , and at optimality no resources will be devoted to system defense.

Interpretation of results. For sufficiently small values of α , it will not be cost-effective to invest resources to defend any components, and the global minimum will be $C_i = 0$ for all i . As α becomes larger, it will eventually become worthwhile to invest in the defense of one or more components. As in the constrained case, since the objective function is not necessarily convex, there can be multiple local minima.

At any local optimum satisfying $C_i > 0$, the marginal effect of additional investment in component i will have a constant value of $-1/\alpha$ (although this need not be true for components that are not defended at optimality). Therefore, as in the constrained case, the optimal allocation of resources for defense depends critically on the cost-effectiveness of defending the various components, as given by the derivatives of the P_i . As in the constrained case, the optimal solution may involve not defending one or more components if they cannot be defended cost-effectively (relative to the value of the system). Unlike in the constrained case, the optimal solution may also involve not defending any components if α is sufficiently small.

Components in Series

Here, we consider a general system with n components in series. In a series system, if even one of the components is disabled, the entire system fails. This can occur for several reasons. In some cases, components may be physically in series (as might be the case, for example, for long pipelines or electricity transmission lines). However, one can also view multiple failure modes for a single component or system as being conceptually in series.

We again assume that the defender wants to achieve either the optimal security within a budget constraint, or the optimal tradeoff between security and the cost of defenses, taking into account the optimal attacker behavior. However, since the components are now in series, the attacker no longer has to disable all components in order to succeed, but instead has a choice of which component(s) to

attack. We assume that the attacker chooses the attack strategy that will maximize the probability of the attack successfully disabling the system. The optimal defense strategy will depend on the attacker's constraints and level of knowledge about the system. Thus, we consider a set of nested optimization problems with varying combinations of assumptions about the attacker and the defender, as described below.

We assume that the attacker is limited to a single attack, and can target only a single component in that attempt. While clearly not fully general, this can be a realistic model in some circumstances (e.g., because of attacker resource constraints, or if a single failed attack would lead to the attacker being detected and disabled). In addition, multiple attacks against a single component can be modeled as a “compound” attack, with a correspondingly greater probability of success. We consider sub-cases in which the attacker has either perfect knowledge of the system's defenses or no knowledge, and in which the defender has a budget constraint or not.

Perfect attacker knowledge of defenses. We assume first that the attacker has perfect knowledge of the system's defenses. This can be a realistic assumption in some cases; e.g., if the attacker has access to inside information. In this case, the optimal attack strategy is to target the weakest component. Thus, the security of the system as a whole is equal to the security of the weakest component. Mathematically, this can be represented by $P(C_1, C_2, \dots, C_n) = \max[P_i(C_i)]$. Here, the defender's resources are best utilized if the $P_i(C_i)$ are brought equal to each other (or as close to equal as possible while satisfying the budget constraint). To see this, consider first the case of two components in series.

Suppose the initial probabilities of attacks succeeding against components 1 and 2 before any defensive resources have been spent satisfy $P_1(0) < P_2(0)$. In this case, since the attacker is assumed to have perfect information about the security of the system, the total probability of a successful attack on the system will be $\max[P_i(0)] = P_2(0)$. Thus, any investment that the defender makes to reduce the success probability of an attack on component 1 will not make any difference to the total security of the system, and hence will not be optimal.

From the above, it is clear that in this case, defensive investments should be allocated so that the success probabilities of attacks against the various components $P_i(C_i)$ are equal to each other (or as close as possible within the budget constraint). Thus, resources should initially be used solely to bring the success probability of an attack against the weaker component to par (or as close to par as possible) with the success probability of an attack against the stronger component—i.e., the component with the lowest initial success probability $P_i(0)$. The extent to which this can (or should) be done will depend on the resource constraint of the defender (or the dollar value α of improving security), since it may be infeasible (or non-optimal) to achieve equality of the success probabilities of attacks against the various components.

To obtain a better understanding about the optimal allocation of defensive investments in this case, we now give a detailed mathematical formulation for optimal defense of a series system against an informed adversary. We consider both constrained and unconstrained defender optimization problems.

Perfect attacker knowledge—constrained case. In this case, the defender has a finite budget C available for strengthening the system, and the objective function to be minimized is $F(C_1, C_2) = \max[P_1(C_1), P_2(C_2)]$. Note that the objective function is monotonically decreasing. Therefore, the minimum will occur where $C_1 + C_2 = C$.

We assume without loss of generality that initially, the probability of an attack against component 1 succeeding is less than that of an attack against component 2; i.e., $P_1(0) < P_2(0)$. Define C_2^* such that $P_1(0) = P_2(C_2^*)$; i.e., $C_2^* = P_2^{-1}[P_1(0)]$. Thus, C_2^* is the level of expenditure required to equalize the success probabilities of attacks against the two components.

If the budget constraint is such that $C \leq C_2^*$, then the optimal strategy for the defender is to use all available resources for improving component 2. In this case, $P_2(C)$ will still be greater than $P_1(0)$ when the budget constraint is reached, so the defender should not strengthen component 1 at all.

If budget constraint C exceeds C_2^* , the defender should first spend an amount C_2^* to strengthen component 2 until it is at par with component 1, and then invest further in both components in such a manner that the success probabilities of attacks against the two components remain at par with each other. At optimality, we will have $P_1(C_1) = P_2(C_2)$. Thus, by the monotonic nature of the objective function, the optimum will occur at the point satisfying $P_1(C_1) = P_2(C - C_1)$, where $C - C_1 > C_2^*$.

Extension to systems with more than two components. As noted above, an attacker with perfect information about system defenses must target the weakest component in order to maximize the probability of an attack succeeding. The defender therefore must again allocate the defensive resources in such a manner that the $P_i(C_i)$ are equal to each other (or as close to each other as possible given the budget constraint).

Because the objective function is monotonically decreasing, we know that the constrained minimum will occur at $C_1 + C_2 + \dots + C_n = C$. The resources of the defender must initially be utilized to try to bring the weakest component—i.e., the component with the highest initial success probability of an attack, $P_i(0)$ —to par with the second weakest component. Then, if that can be achieved, those two components can be strengthened to bring them to par with the third weakest (or as close as possible), and so on until the resource constraint is reached. If the defender is able to bring all components to par with the strongest component, any remaining resources should then be allocated in such a manner that the $P_i(C_i)$ remain equal.

Interpretation of results. The above discussion indicates that if the attacker knows about the system's defenses, the defender's options for protecting a series system are limited. In particular, the attacker's ability to respond strategically to the defender's investments deprives the defender of the ability to allocate defensive investments according to their cost-effectiveness (as measured, for example, by the slopes of the P_i). Thus, when components are in series, and potential attackers know (or can easily learn) about the effectiveness of any defensive measures, defensive investments must essentially equalize the strength of all components in order to be beneficial.

Dresher [1961] earlier noted in the military context that, for optimal allocation of defensive resources, "It is necessary that each of the defended targets yield the same payoff to the attacker." Thus, even if some components can be hardened

inexpensively, focusing protective investments on only those will lead to wasted resources if adversaries can choose to attack components that are more costly to harden. This suggests that defense will generally be more costly when the adversary knows about the system defenses. Similarly, in an analysis of asymmetric warfare against civilian targets, Ravid [2001] argues that when the adversary can change targets in response to defensive investments, “investment in defensive measures, unlike investment in safety measures, saves a lower number of lives (or other sort of damages) than the apparent direct contribution of those measures.” He concludes that security improvements will tend to be more costly than safety improvements.

Perfect attacker knowledge—unconstrained case. Here, the defender has no budget constraint. As before, the constant α expresses the dollar value of reducing the success probability of an attack against the system from one to zero, and we assume that the defender wishes to optimize the tradeoff between security and cost allocation. In this case, the probability that an attack against the system will succeed remains $P(C_1, C_2) = \max[P_1(C_1), P_2(C_2)]$. However, the objective function now takes into account the tradeoff α and the costs C_i :

$$F(C_1, C_2) = \alpha \max [P_1(C_1), P_2(C_2)] + C_1 + C_2 \quad (7)$$

This function is convex. Thus, any local minimum must be a global minimum. However, the optimal solution can take on several different forms. For sufficiently small values of α , it will not be cost-effective to invest resources to defend either of the components, and the global minimum will occur at $C_1 = C_2 = 0$. As α becomes larger, it will eventually become worthwhile to invest in defense.

If the global minimum has both $C_i > 0$, then we must have $P_1(C_1) = P_2(C_2)$; i.e., $C_2 = P_2^{-1}[P_1(C_1)]$. Substituting this expression into Equation (7), we can now rewrite the objective function in the form $F\{C_1, P_2^{-1}[P_1(C_1)]\} = \alpha P_1(C_1) + C_1 + P_2^{-1}[P_1(C_1)]$. While the original objective function was not differentiable, the revised objective function is differentiable for all non-negative C_1 such that $P_2^{-1}[P_1(C_1)]$ is greater than zero. Moreover, since the objective function is convex, differentiating with respect to C_1 and setting the result equal to zero yields a global minimum satisfying:

$$\begin{aligned} \frac{F\{C_1, P_2^{-1}[P_1(C_1)]\}}{\partial C_1} &= \alpha \frac{\partial P_1(C_1)}{\partial C_1} + 1 + \frac{\partial P_2^{-1}[P_1(C_1)]}{\partial P_1(C_1)} \frac{\partial P_1(C_1)}{\partial C_1} \\ &= \alpha P_1'(C_1) + 1 + P_2^{-1}[P_1(C_1)] P_1'(C_1) = 0 \end{aligned}$$

Therefore, if both of the C_i are positive at optimality, then the optimal solution will satisfy $P_1'(C_1) = -1/\{\alpha + P_2^{-1}[P_1(C_1)] P_1'(C_1)\}$.

Special case. Assume that the $P_i(C_i)$ are identical; i.e., $P_1(C_1) = P_2(C_1)$. Thus, we can now write the objective function as $F(C_1, C_2) = \alpha \max[P_1(C_1), P_1(C_2)] + C_1 + C_2$. If α is sufficiently large to justify some non-zero level of defensive investment, then the optimum will occur when $P_1(C_1) = P_1(C_2)$; i.e., when $C_1 = C_2$. In this case, rewriting the objective function in terms of C_1 alone yields $F\{C_1, P_1^{-1}[P_1(C_1)]\} = \alpha P_1(C_1) + 2$

C_1 . Differentiating and setting the result to zero, we find that the optimum will occur at the point satisfying $P_i'(C_i) = -2/\alpha$.

Extension to systems with more than two components. We have assumed that the attacker possesses perfect information regarding the system defenses, and will always target the weakest component. Therefore, regardless of the number of components, the defender should allocate the defensive resources in such a manner that the $P_i(C_i)$ are brought equal to each other, or as near to equal as can be justified given the value of α . As in the two-component case, the cost-effectiveness of strengthening the various components will affect the total of the defensive investments, but not the allocation of that total investment among the various components.

Interpretation of results. Unlike in the parallel case, the optimal defensive strategy does not depend strongly on the cost-effectiveness of strengthening the various components. While the optimal solution does depend on the derivatives P_i' (and also the derivatives of the P_i^{-1}), those derivatives now influence only the total level of investment $C_1 + C_2$, since beyond some point, further increases in defensive investments will no longer be justified. Thus, the magnitude of the total defensive investment depends on the value of α , with small values of α leading to small or even zero total investment. However, the goal of bringing the weaker components to par with the stronger ones will completely dictate the allocation of any non-zero investment among the various components.

No attacker knowledge of defenses. When the attacker has no information about the system's defenses, we assume that the attacker will target one component at random, without regard to the defensive investments C_i . Consider a system with two components, and suppose that any attack will target component 1 with probability q and component 2 with probability $1 - q$. The success probability of an attack will be $P(C_1, C_2) = q P_1(C_1) + (1 - q) P_2(C_2)$. $P(C_1, C_2)$ is clearly convex, since $P_1(C_1)$ and $P_2(C_2)$ are convex, and q and $1 - q$ are non-negative. As before, we consider both constrained and unconstrained cases.

No attacker knowledge—constrained case. Here, the defender is assumed to have a resource constraint C , and the objective function to be minimized by the defender is $F(C_1, C_2) = P(C_1, C_2) = q P_1(C_1) + (1 - q) P_2(C_2)$. This function is monotonically decreasing, so the optimum will occur at $C_1 + C_2 = C$. To find the minimum, we substitute $C_2 = C - C_1$ into the objective function, differentiate with respect to C_1 , and equate the result to zero, yielding:

$$\frac{\partial F(C_1, C - C_1)}{\partial C_1} = q P_1'(C_1) - (1 - q) P_2'(C - C_1) = 0$$

By convexity, any local minimum must be the unique global minimum. Moreover, if the local minimum occurs for C_1 between zero and C , it will satisfy $q P_1'(C_1) = (1 - q) P_2'(C - C_1)$. The optimum can also occur with C_1 equal to zero if component 2 can be

defended much more cost-effectively than component 1, or with C_1 equal to C if the converse, in which case it need not satisfy this first-order condition.

Special case. Let the conditional probability that each component is attacked be 0.5. To optimize the system defenses, the defender should achieve $P_1'(C_1) = P_2'(C - C_1)$ if this can be achieved for $0 < C_1 < C$. At this point, the improvement in the security of component 1 by increasing C_1 will be exactly equal to the corresponding deterioration in the security of component 2. Thus, when the conditional probabilities of attacks on the two components are equal, the defender should invest more in the component that is more cost-effective to strengthen. Moreover, if the functions P_i are identical, then the defender should spend the same level of resources on each component.

Extension to systems with more than two components. For a system with n components, the objective function that the defender wishes to minimize is:

$$F(C_1, C_2, \dots, C_n) = \sum_{i=1}^n q_i P_i(C_i)$$

where q_i is the probability that an attack will target component i , and the q_i are assumed to sum to one. The global minimum will occur at $C_1 + C_2 + \dots + C_n = C$, or in other words $C_1 = C - (C_2 + \dots + C_n)$. Substituting this value of C_1 into the objective function, we get:

$$F(C_1, C_2, \dots, C_n) = q_1 P_1[C - (C_2 + \dots + C_n)] + \sum_{i=2}^n q_i P_i(C_i)$$

To find the global optimum, we differentiate the objective function with respect to C_i (for $i \neq 1$) and equate the result to zero, yielding:

$$\frac{\partial F(C_1, C_2, \dots, C_n)}{\partial C_i} = q_1 P_1'[C - (C_2 + \dots + C_n)] + q_i P_i'(C_i) = 0$$

At optimality, we will have $q_1 P_1'[C - (C_2 + \dots + C_n)] = q_i P_i'(C_i)$, if this occurs for positive values of the C_i . Thus, the marginal effects of additional resources spent on all defended components will be equal; i.e., components with C_i non-zero will all have the same marginal effectiveness for additional investment. This is similar to the result for a two-component system. The level of resources spent on component i will also generally increase with the conditional probability of attack q_i .

Interpretation of results. The above results show that the defender has much more flexibility to allocate resources between the components of a series system cost-effectively when the attacker has no knowledge of the system defenses than when the attacker has perfect information about the defenses. In other words, for a fixed budget, the defender will generally be able to achieve better system security when the attacker has little or no knowledge of the system defenses. This demonstrates the potential importance of secrecy (and/or deception) as defensive strategies.

As one example, the proposal to sterilize mail to protect against future anthrax attacks [Florig 2002] might be an effective defense if the installation of such sterilization equipment were a secret. However, this would probably not be possible in our open society, given the outcry of public concern about the threat of anthrax spores sent through the mail. If the installation of anthrax sterilization equipment were public knowledge, it would likely only cause future attackers to find a means of delivery other than the public mail. In that case, the proposed large investment in sterilization equipment may never sterilize a single anthrax spore!

No Attacker Knowledge—Unconstrained Case. Here, the defender has no budget constraint, but the parameter α gives the importance assigned to system security relative to the cost of defensive investments. The probability of an attack succeeding remains $P(C_1, C_2) = q P_1(C_1) + (1 - q) P_2(C_2)$. However, the objective function must now take into account the tradeoff α and the costs C_i , yielding $F(C_1, C_2) = \alpha [q P_1(C_1) + (1 - q) P_2(C_2)] + C_1 + C_2$. The objective function $F(C_1, C_2)$ is still convex, since α is non-negative and the cost terms are linear. To find the global minimum for this function, we differentiate with respect to C_1 and C_2 , respectively, and equate the first derivatives to zero:

$$\frac{\partial F(C_1, C_2)}{\partial C_1} = \alpha q P_1'(C_1) + 1 = 0$$

$$\frac{\partial F(C_1, C_2)}{\partial C_2} = \alpha (1 - q) P_2'(C_2) + 1 = 0$$

By the convexity of the objective function, if the global minimum occurs where C_1 —respectively, C_2 —is positive, then at optimality, we will have $P_1'(C_1) = -1/(\alpha q)$ —respectively, $P_2'(C_2) = -1/[\alpha (1 - q)]$. However, for sufficiently small values of α , the global minimum may have one or both of the C_i equal to zero, in which case the above condition need not be satisfied for the variable(s) equal to zero at optimality.

Special cases.

1. As in the constrained case above, if the conditional probability that each component is attacked equals 0.5, then the defender should invest more resources in the component that is more cost-effective to strengthen.
2. When the functions $P_i(C_i)$ are identical, the defender should invest money in the component that the attacker is more likely to target, as is intuitive.

Extension to systems with more than two components. As above, for any component i receiving a non-zero defensive investment at optimality, we will have $P_i'(C_i) = -1/(\alpha q_i)$, where q_i is the conditional probability of an attack on component i . Thus, the results are similar to those for two components.

Interpretation of results. Increases in the value of α will lead to increases in the optimum values of the C_i , at least for those components that are already receiving

positive levels of investment at optimality. In addition, an increase in the conditional probability of an attack on a particular component will generally increase the optimum investment in that component. Also, in this case, the optimal levels of investment in the various components can be determined independently of each other, since there is no longer a budget constraint to induce coupling between them.

Conclusions and Future Work

Our results suggest that defending series systems against informed attackers is an extremely difficult challenge. This emphasizes the importance of redundancy as a defensive strategy—especially if attacks against the redundant components are also likely to succeed or fail approximately independently of each other (e.g., due to factors such as spatial separation and functional diversity). Our results also support the idea that secrecy or even deception can be an important strategy for improving security (especially for series systems), and/or reducing defensive costs.

Obviously, it is important to extend this work to more general systems, rather than only series and parallel architectures. Unfortunately, finding optimal attack strategies for arbitrary systems is NP-hard, as can be shown from results in Cox et al. [1989, 1996]. Therefore, we plan to adapt near-optimal heuristics for least-cost diagnosis from Cox et al. to derive near-optimal heuristic attack strategies, and identify optimal or near-optimal defenses against such attacks. We also hope to determine when those heuristic attack strategies will be optimal.

Considering more general system structures (e.g., using graph theory methods) would allow us to model attacks for instrumental purposes, such as internet “worms,” in which poorly defended computers are compromised in order to use them in an attack against a more valuable but better defended system. One could also use similar approaches to investigate quantitatively the relative merits and limitations of perimeter defense in computer networks; see, for example, Wulf and Jones [2002].

In addition, we plan to extend our models to consider systems in which the various components have different inherent “values” to the defender (and hence to the attacker), rather than being valuable only insofar as they contribute to the operability of the system as a whole. For example, a defender might be more concerned about an attacker successfully accessing a database of critical information than about an attack that merely renders a computer system unusable. Such models would help us investigate the appropriateness of the heuristic suggested by O’Hanlon et al. [2002], of allocating resources to protect only the most valuable assets. Preliminary results suggest that the suitability of this heuristic will depend to some extent on the assumptions made about attacker goals and motivations. For example, the proposed heuristic may be optimal if attackers choose targets solely according to their value, but not if they also consider the probability of success in their choice of targets.

Also, it would be worthwhile to extend our models to include a time dimension, rather than the current static or “snapshot” view of system security. This would allow us to model imperfect attacker information (including Bayesian updating of the probability that an attack will be successful, as estimated by both attacker and defender), and the possibility of multiple attacks (either attacks against multiple components, or multiple attacks against the same component).

One could also apply similar methods to model attacks involving opportunistic rather than determined adversaries. In this case, it may be helpful to formulate the problem as a game between multiple defenders [Kunreuther and Heal 2002], rather than a game between a single defender and a single adversary.

Acknowledgments

This material is based upon work supported in part by the US Army Research Laboratory and the US Army Research Office under grant number DAAD19-01-1-0502. Any opinions, findings, and conclusions or recommendations expressed in this document are those of the authors and may reflect the views of the sponsors.

References

- Cox, L.A., Y. Qiu, and W. Kuehner (1989). "Heuristic least-cost computation of discrete classification functions with uncertain argument values." *Ann. of Operations Research*, 21, 1-30.
- Cox, L.A., S. Chiu, and X. Sun (1996). "Least-cost failure diagnosis in uncertain reliability systems." *Reliability Eng. and System Safety*, 54(2-3), 203-216.
- Dresher, M. (1961). *Games of Strategy: Theory and Applications*. Englewood Cliffs, NJ: Prentice-Hall.
- Florig, H.K. (2002). "Is safe mail worth the price?" *Science*, Feb. 22, 1467-1468.
- Kunreuther, H. and G. Heal (2002). "Interdependent security: the case of identical agents." Insurance Project Workshop, National Bureau of Economic Research, Inc., Cambridge, MA, <http://www.nber.org/~confer/2002/insw02/kunreuther.pdf>.
- O'Hanlon, M., P. Orszag, I. Daalder, M. Destler, D. Gunter, R. Litan, and J. Steinberg (2002). *Protecting the American Homeland*. Washington, DC: Brookings Institution.
- President's Commission on Critical Infrastructure Protection (1997). *Critical Foundations: Protecting America's Infrastructures*. Washington, DC, <http://www.terrorism.com/homeland/pccipreport.pdf>.
- Ravid, I. (2001). "Theater ballistic missiles and asymmetric war." Draft from author.
- Schneier, B. (2001). "Military history and network security" (sidebar). Counterpane Internet Security, Cupertino, CA, <http://www.counterpane.com/msm.pdf>.

Schudel, G. and B. Wood (2000). "Modeling behavior of the cyber-terrorist." In Anderson, R.H., T. Bozek, T. Longstaff, W. Meitzler, M. Skroch, and K. Van Wyk, *Conference Proceedings: Research on Mitigating the Insider Threat to Information Systems—#2*. Santa Monica, CA: Rand.

Wulf, W.A. and A.K. Jones (2002). "Cybersecurity." *The Bridge*, 32(1), 41-45.