**SANDIA REPORT**
SAND2006-4155
Unlimited Release
Printed July 2006

# Physical Security and Vulnerability Modeling for Infrastructure Facilities

Dean A. Jones, Chad E. Davis, Mark A. Turnquist, and Linda K. Nozick

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia is a multiprogram laboratory operated by Sandia Corporation,
a Lockheed Martin Company, for the United States Department of Energy's
National Nuclear Security Administration under Contract DE-AC04-94AL85000.

Sandia National Laboratories

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

**NOTICE:** This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government, nor any agency thereof, nor any of their employees, nor any of their contractors, subcontractors, or their employees, make any warranty, express or implied, or assume any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represent that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government, any agency thereof, or any of their contractors or subcontractors. The views and opinions expressed herein do not necessarily state or reflect those of the United States Government, any agency thereof, or any of their contractors.

Printed in the United States of America. This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

| | |
|---|---|
| Telephone: | (865) 576-8401 |
| Facsimile: | (865) 576-5728 |
| E-Mail: | reports@adonis.osti.gov |
| Online ordering: | http://www.osti.gov/bridge |

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Rd.
Springfield, VA 22161

| | |
|---|---|
| Telephone: | (800) 553-6847 |
| Facsimile: | (703) 605-6900 |
| E-Mail: | orders@ntis.fedworld.gov |
| Online order: | http://www.ntis.gov/help/ordermethods.asp?loc=7-4-0#online |

SAND2006-4155
Unlimited Release
Printed July 2006

# Physical Security and Vulnerability Modeling for Infrastructure Facilities

Dean A. Jones, Chad E. Davis, Mark A. Turnquist, and Linda K. Nozick
Knowledge Systems Design and Engineering, Department 6223

Sandia National Laboratories
P.O. Box 5800
Albuquerque, NM 87185-1138

## Abstract

A model of malicious intrusions in infrastructure facilities is developed, using a network representation of the system structure together with Markov models of intruder progress and strategy. This structure provides an explicit mechanism to estimate the probability of successful breaches of physical security, and to evaluate potential improvements. Simulation is used to analyze varying levels of imperfect information on the part of the intruders in planning their attacks. An example of an intruder attempting to place an explosive device on an airplane at an airport gate illustrates the structure and potential application of the model.

# Contents

1. Introduction .......... 7
2. Intrusion Attempts at a Node .......... 9
3. Expanding to the System Level .......... 13
4. An illustrative application .......... 15
5. Representing Imperfect Information .......... 19
6. Illustrative Simulation Results .......... 21
7. Optimizing Resource Allocation for Security Improvement .......... 25
8. Extensions .......... 27
9. Conclusions .......... 29
10. References .......... 31

# Figures

Figure 1. A Hidden Markov Model characterizing an attack at a system node. .......... 9
Figure 2. Aggregated abstraction of the HMM at a node .......... 10
Figure 3. Simple system-level network. .......... 13
Figure 4. Illustrative network for analyzing an attempted placement .......... 15
Figure 5. Summary of intruder strategy and probability of success. .......... 17

# Tables

Table 1: Example data for network nodes. .......... 16
Table 2: Probability of detection for possible moves. .......... 16
Table 3: Summary of results when probability estimates are unbiased and coefficient of variation is 0.1 .......... 21
Table 4: Summary of results when probability estimates are unbiased and coefficient of variation is 0.3. .......... 22
Table 5: Summary of results when probability estimates are biased on arc 4-6 and 4-7. .......... 23

# 1. Introduction

There is widespread interest in protection of critical infrastructures from malicious attack. The attacks might be either physical intrusions (e.g., to steal vital material, plant a bomb, etc.) or cyber intrusions (e.g., to disrupt information systems, steal data, etc.). The attackers may be international terrorists, home-grown hackers, or ordinary criminals. In 1997, the report of the U.S. President's Commission on Critical Infrastructure Protection identified eight critical infrastructures "whose incapacity or destruction would have a debilitating impact on our defense and economic security" [11]. In subsequent years, this list of critical infrastructures was expanded and a set of 13 critical infrastructure sectors are included in the National Strategy for Homeland Security [4]. These 13 are: agriculture, food processing, water, public health, government, emergency services, banking and finance, telecommunications, energy, transportation, the chemical industry, postal and shipping services, and the defense industrial base.

In this analysis, we focus primarily on transportation facilities, but the approach we suggest could also be used in other infrastructure contexts. For example, a similar type of analysis has been applied to information systems [3]. The objective of the analysis presented here is to provide guidance to system owners and operators regarding effective ways to reduce vulnerabilities of specific facilities. To accomplish this, we develop a Markov Decision Process (MDP) model of how an intruder might try to penetrate the various barriers designed to protect the facility. This intruder model provides the basis for consideration of possible strategies to reduce the probability of a successful attack on the facility.

We represent the system of interest as a network of nodes and arcs. Nodes represent barriers that an intruder must penetrate, and arcs represent movements between barriers that an intruder can make within the system. The adversaries first must penetrate entry points to the system, and if an attempted penetration at a particular entry node is successful, they can traverse edges from the successfully breached node to other nodes in the network that are connected to the one breached. Traversing an edge entails a risk of detection. The adversary is assumed to make the decision that maximizes the probability of successful attack.

Several previous authors have used graph-based methods to represent attackers or defenders in security analyses. Phillips and Swiler [10] introduced the concept of an "attack graph" to represent sets of system states and paths for an attacker to pursue an objective in disrupting an information system. Several subsequent papers (e.g., [5], [14], [16]) have extended these initial ideas.

A number of authors have used Markov models to represent uncertainties in system state in the face of attacks, especially in computer systems (e.g., [5], [7], [14], [15]). In particular, Hidden Markov Models (HMM) focus on intruder detection using indicators that indirectly reflect potential attacker activities (see, for example, [8], [15], [17]).

Jha et al. [5] introduced the idea of using Markov Decision Processes (MDP) for situations in which the intruder's path is probabilistic. By interpreting attack graphs as Markov Decision Processes they computed a probability of intruder success for each attack represented by the graph. In the current work, we also use the idea of computing the probability of a successful attack by characterizing the problem as an MDP. However, our graph structure is different from the normal attack graph structure used in information systems, and thus the underlying network over which the MDP is formulated is different from that used in [5].

Our primary attention is on a class of adversaries that is rational – i.e., the adversaries follow a strategy that maximizes the probability of their attack being successful.  Their ability to actually optimize their attacks depends also on the level of information that they have about the system, and part of our analysis is to focus on how varying levels of information about the infrastructure system affects the strategies of potential intruders, how the overall probability of intruder success is affected by their level of information, and what implications this has for effective defense of the system against intrusion.

We first construct an HMM to represent an intruder's actions at a single node (barrier) in a system. Then we develop an aggregated representation of that single-node model for inclusion in an MDP model of intruder strategy within a network representation of the entire system. Finally, we explore how varying levels of information on the part of the intruder affect the intruder's strategy and likelihood of success.

# 2. Intrusion Attempts at a Node

An attempt to penetrate a system barrier (node) and the interaction between the intruder and the intrusion-detection system is modeled using a Hidden Markov Model (HMM). The general concept of such a model is represented in Figure 1. The intruder's actions (the lower portion of the diagram) are assumed to progress through a set of states as a Markov process. The diagram in Figure 1 shows a simplified representation in which transitions are only to sequential states, but the transition matrix used can be more general. Occupancy of various states may result in emanations that are observable by the system operator (represented by the "signals" in Figure 1). For example, the intruder may be attempting to pick the lock of a door where there is video surveillance. Picking the lock requires an uncertain amount of time, represented by transition through a series of Markov states. While the intruder occupies those states (i.e., during the time that the intruder is attempting to pick the lock), there is a probability that his/her presence will be detected by the video surveillance system. The general structure of the HMM allows considerable flexibility in defining various types of signals and resulting actions by the system operator. For example, some signals may cause an increased level of surveillance without an alarm being raised. For our current purposes, we use a straightforward definition that a recognized signal from any state constitutes detection and the end of the attempted intrusion. If the intruder reaches a "breach" state without being detected, we say that the node (barrier) has been breached, and no further emanations will cause the system to detect the intruder at that node. We also include a "retreat" state that corresponds to an unsuccessful, but undetected, attempt to penetrate the barrier. In that outcome, the intruder can withdraw without raising an alarm.

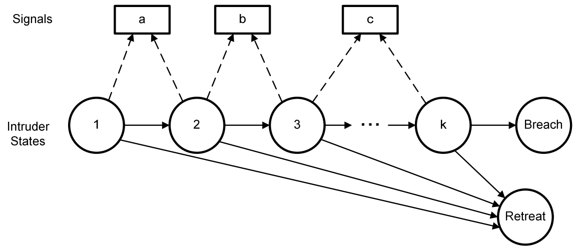

**Figure 1. A Hidden Markov Model characterizing an attack at a system node.**

We use a discrete-time, discrete-state HMM characterized by the following equations:

$$X_{n+1} = A^T X_n \tag{1}$$

$$Y_n = BX_n \tag{2}$$

for transition steps $n = 1, 2, \ldots, \infty$. The state of the system (i.e., presence of the intruder in some node in the lower portion of Figure 1) is represented by the (column) probability vector, $X$. The dynamics of the system are governed by (1), where $A$ is a transition matrix (i.e., it satisfies the properties $a_{ij} \geq 0$ and $\sum_j a_{ij} = 1$.) The states of the system are not observed directly. The process $Y$ is observed, which is a function of the state of the underlying Markov process, $X$. Each column of $B$ specifies a conditional probability distribution over the possible observations, given that the underlying (hidden) system is in a particular state. The estimated values for B in a given application should reflect any efforts that might be taken by an intruder to reduce the likelihood of detection (e.g., attempting to defeat sensors, create diversions, etc.).

For our purposes, we assume that *A* and *B* are known (or have been estimated). We want to use the estimated HMMs at various nodes as the basis for a network-level model of intruder strategy. In large networks, it is useful to abstract the HMM at node *v* to a simpler representation, as shown in Figure 2. An intruder enters an "Attempt" state for that barrier (node). The intruder continues to occupy that state until the attempted penetration is detected (and an alarm is raised), the penetration is successful and the barrier is breached, or the intruder retreats.

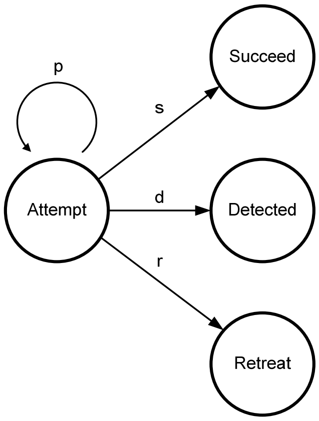

**Figure 2. Aggregated abstraction of the HMM at a node.**

To make the abstraction in Figure 2 useful, we must be able to derive the transition probabilities *p*, *s*, *d* and *r* from the underlying *A* and *B* matrices of the HMM. The "attempt" state in Figure 2 is a transient state, and the other three states are absorbing. The transition probabilities *s, d* and *r* are specified so that the probabilities of detection, successful breach and retreat match those from the original HMM. The transition probability *p* is specified so the expected length of residence in the "attempt" state matches the duration of the attempted penetration in the original HMM.

To determine the transition probabilities, we construct an augmented state space for the HMM by adding a detection state. The transition probabilities to the detection state are given by the $b_{ij}$ values from the *B* matrix corresponding to emanations that are specified to cause detection. The original transition probabilities (in the A matrix) are adjusted to account for the probability of detection. The resulting transition matrix for the augmented state space will be denoted as *P*.

$$P = \begin{bmatrix} Q & Z \\ 0 & H \end{bmatrix} \tag{3}$$

The submatrix *Q* represents transitions among the transient states, *Z* represents transitions from the transient states into the absorbing states, and *H* represents transitions within the set of absorbing states. In applications of interest here, *H* is an identity matrix.

The Fundamental Matrix $\Phi = (I - Q)^{-1}$ contains elements $\phi_{ij}$, interpreted as the expected number of visits to state *j* before absorption, given that the system started in state *i* (see, for example, [13]). In general, the expected number of transitions prior to absorption for an attacker who enters in state *i* is:

$$n_i = \sum_j \phi_{ij} \tag{4}$$

For the reduced representation in Figure 2, the expected number of transitions prior to absorption is:

$$n = \frac{1}{1-p} \tag{5}$$

If there is only one entry state, *i*, in the original HMM, then we can equate (4) and (5) to solve for *p* as:

$$p = 1 - \frac{1}{\sum_j \phi_{ij}} \tag{6}$$

If the HMM has multiple entry states, denoted by a set *E*, with probability of entry in state $i \in E$ denoted by $\rho_i$, we can compute the unconditional expected number of transitions prior to absorption as $\overline{n} = \sum_{i \in E} \rho_i n_i$, and the equation for *p* becomes:

$$p = 1 - \frac{1}{\overline{n}} \tag{6'}$$

In Markov chains that have both transient and absorbing states, if *j* is one of the absorbing states, the probability that the system is absorbed in state *j*, given that the initial state was state *i*, is given by the *ij*th element of the matrix ΦZ (for a proof of this result, see [2], page 157). We denote this conditional probability as $f_i(j)$:

$$f_i(j) = [\Phi Z]_{ij} \tag{7}$$

Thus, the probability of successful breach in the original HMM is:

$$\Pr(success) = \sum_{i \in E} f_i(success) \tag{8a}$$

Similarly, the probabilities of detection and retreat are:

$$\Pr(detected) = \sum_{i \in E} f_i(detected) \tag{8b}$$

$$\Pr(retreat) = \sum_{i \in E} f_i(retreat) \tag{8c}$$

In the reduced state representation (Figure 2), equation (7) allows us to write parallel expressions:

$$\Pr(success) = \frac{s}{s+d+r} \tag{9a}$$

$$\Pr(detected) = \frac{d}{s+d+r} \tag{9b}$$

$$\Pr(retreat) = \frac{r}{s+d+r} \tag{9c}$$

By equating (8a) and (9a), (8b) and (9b), and (8c) and (9c), and by noting that $s+d+r=1-p$, we can solve for the required transition probabilities *s*, *d* and *r* in the reduced representation.

The value of the aggregated representation is that it allows us to construct a Markov Decision Process (MDP) of the intruder's strategy at the system level, without carrying along all the detail of states within each node. This is the focus of the following section.

# 3. Expanding to the System Level

At the system level, we represent a network of barriers and potential movements as shown in the simple example in Figure 3. Each node can be expanded using a representation like the one in Figure 2. If the intruder is successful at breaching a particular barrier, he/she has choices about where to go next (which arc to cross). Crossing arc *ij* entails a probability of detection $\delta_{ij}$ , and this is represented in the transition matrix.

We can pose the problem of finding the intruder's optimal strategy as an MDP over an infinite horizon. We define the expected "reward" to the intruder as a value associated with reaching the "success" state of a goal node (such as node 8 in the example in Figure 3), which represents an undetected exit from the system after accomplishing a desired action (such as placing a bomb, etc.). If we define this reward value as 1, then the expected rewards calculated at all earlier nodes in the network can be interpreted as probabilities of success, given that the intruder has reached that node.

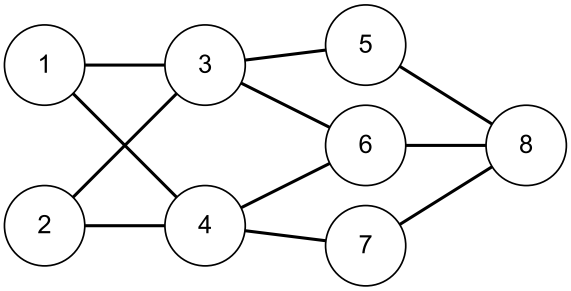

**Figure 3. Simple system-level network.**

We assume that the objective of the intruder is to maximize his/her expected reward (probability of successful attack), and we examine the problem of finding the optimal strategy for this objective. Solving this problem positions us to adopt the perspective of the system operator and consider the actions that can have the largest impact on reducing the probability of successful intrusions.

If the intruder is in state *i* and chooses action $a_i$, we denote the expected value of the future stream of rewards by $w(i,a_i)$. Each possible action $a_i$ implies a change in the transition probabilities that govern the process. We denote the elements of the transition matrix resulting from choosing action $a_i$ as $P_{ij}(a_i)$. The MDP we define for this problem is positive bounded, and we can find the optimal policy through either policy iteration or linear programming.

From a computational standpoint, policy iteration is generally preferable to linear programming for finding solutions, but the linear programming formulation can yield insights that are significant for our current purposes. Puterman [12] describes the linear programming formulation for positive bounded expected total reward models. The formulation seeks the decision policy (choice of $a_i$) that maximizes the expected value of the reward stream, $w(i,a_i)$. We denote the resulting optimal expected value as $w^*(i)$.

As [12] describes in detail, the set of $w^*(i)$ is the smallest set of values of $w(i)$ for which the following inequalities hold for all states, *i*:

$$w(i) \geq R_i(a_i) + \sum_j P_{ij}(a_i) w(j) \tag{10}$$

where $R_i(a_i)$ is the immediate reward for selecting action $a_i$ when the system state is $i$. In our application, $R_i(a_i) = 0$ for all states $i$ other than the goal state, $g$, and $R_g(a_g) = 1$ for the dummy action, $a_g$, after achieving the goal state.
If we then introduce an arbitrary set of positive scalars, $\beta_i$, with the requirement that $\sum_i \beta_i = 1$, the linear program can be written as follows:

$$\min \sum_i \beta_i w(i) \tag{11}$$

subject to:

$$w(i) - \sum_j P_{ij}(a_i) w(j) \geq R_i(a_i) \qquad \forall\, i, a_i \tag{12}$$

$$w(i) \geq 0 \qquad \forall\, i \tag{13}$$

This linear program has a dual that can be expressed as follows:

$$\max \sum_i \sum_{a_i} R_i(a_i) x_i(a_i) \tag{14}$$

subject to:

$$\sum_{a_i} x_i(a_i) - \sum_j \sum_{a_i} P_{ij}(a_i) x_i(a_i) \leq \beta_i \qquad \forall i \tag{15}$$

$$x_i(a_i) \geq 0 \quad \forall i, a_i \tag{16}$$

In our case, because all but one of the $R_i(a_i)$ values are zero, the dual objective function can be simplified to:

$$\max \quad x_g(a_g) \tag{16'}$$

The primal linear program has many more constraints than variables, so it is more effective to solve the dual problem. In addition, it can be shown (see [12]) that in an optimal solution to the dual problem (14)–(16), there is no more than one non-zero $x_i(a_i)$ for each state $i$. The $a_i$ for which $x_i(a_i)$ is non-zero indicates the optimal action $a_i^*$ for each $i$. The shadow prices on the dual constraints (15) are the values of $w^*(i)$, indicating the probability of successful attack, given that the intruder has reached state $i$.

# 4. An illustrative application

As an example of system-level analysis for a specific infrastructure facility, consider an intruder who is attempting to place an explosive device aboard an aircraft while it is sitting at an airport gate, with the intent that it will explode later after the aircraft is in flight. A simplified representation of the barrier network and possible intruder actions is shown in Figure 4 (the network structure is the same as in Figure 3, but the nodes and links have now been labeled as specific barriers and movements).

The intruder must first gain access to the apron area of the terminal. We postulate that this can occur either by gaining illicit access through the employee gate (e.g., by stealing an employee ID and using it to enter the area), or by entering in a service vehicle at a gate (e.g., in a catering truck). If the intruder is successful in getting access to the area, he/she must then impersonate a legitimate worker in the aircraft gate area – either an airline employee or a service contractor. The "cross-over" arcs between "entry" and "impersonation" in Figure 4 indicate that even if the intruder gains access to the apron area using an employee ID, he/she may switch ID's and impersonate a service contractor within the area (or vice versa). This impersonation must be successful for the period of time required to get from the entrance to the aircraft itself.

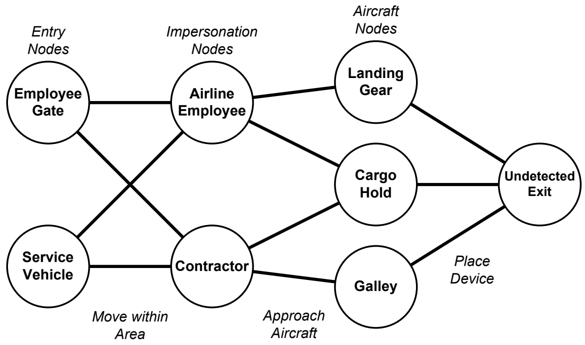

**Figure 4. Illustrative network for analyzing an attempted placement of an explosive device on an aircraft.**

Approaching the aircraft carries a risk of detection, and the approachable areas on the aircraft if the intruder is impersonating an employee may be different from those that are approachable if he/she is impersonating a service contractor. For example, a person who appears to be an airline maintenance employee might not attract attention approaching the under-wing area around the landing gear, whereas a person who appears to be a catering contractor would. For purposes of this example, we consider in Figure 4 three areas of the aircraft where an explosive device might be hidden – inside the wing around the landing gear, in the cargo hold, or in the catering supplies delivered to the galley.

If access to the aircraft is gained, the device must be placed without arousing suspicion. This is represented by the arcs connecting the aircraft area nodes to the exit node. Each of these arcs has a probability of detection.

Finally, if the intruder succeeds in gaining access to the aircraft and placing the device, he/she must exit without detection, and this represents the last barrier. Our modeling premise is that if the intruder is detected after placing the device, it will trigger a thorough search of the aircraft and the device will be discovered, so that the attempted attack will be foiled.

Table 1 summarizes the hypothetical node data used for the example analysis, and Table 2 shows the probabilities of detection used for the arcs in the example network. Note that we assume there is no retreat at the stage of exiting after placing the device – at that stage either the attack is successful or it is detected. Also note that the probability of detection on the arcs leading to the "impersonation" nodes is zero. This is because we are treating impersonation process (and time) as a barrier (node), so the probability of detection is lumped at the nodes, rather than on the arcs.

If an intruder knew the structure of the network (Figure 4) and the values in Tables 1 and 2, we would consider him/her to be perfectly informed. Under this assumption, an optimal intrusion strategy (i.e., one that maximizes the probability of successful attack) can be constructed by solving the MDP. For the set of input data in Figure 4 and Tables 1 and 2, the solution for the optimal intruder strategy can be summarized as shown in Figure 5. To the left of each node is the probability of successful attack, given that the intruder is "arriving at" that barrier. To the right of each node is the probability of success, given that the intruder has successfully negotiated that barrier. There is only one value shown for the exit node (i.e., the "approaching" probability), because once that node is successfully negotiated, the attack has been a success, by definition.

**Table 1. Example data for network nodes.**

| Node (see Figure 4) | Expected Time for Attempted Breach (min) | Prob. of Success | Prob. of Detection | Prob. of Retreat |
|---|---|---|---|---|
| Employee Gate | 1 | 0.2 | 0.65 | 0.15 |
| Service Gate | 2 | 0.25 | 0.7 | 0.05 |
| Impersonate Employee | 10 | 0.2 | 0.6 | 0.2 |
| Impersonate Contractor | 15 | 0.4 | 0.5 | 0.1 |
| Landing Gear | 5 | 0.15 | 0.8 | 0.05 |
| Cargo Hold | 3 | 0.1 | 0.75 | 0.15 |
| Galley | 15 | 0.15 | 0.75 | 0.1 |
| Undetected Exit | 10 | 0.8 | 0.2 | 0 |

**Table 2. Probability of detection for possible moves.**

| Arc | Prob. of Detection |
|---|---|
| Empl. Gate – Impersonate Employee | 0 |
| Empl. Gate – Impersonate Contractor | 0 |
| Service Gate – Impersonate Empl. | 0 |
| Service Gate – Impersonate Contr. | 0 |
| Impersonate Empl. – Landing Gear | 0.7 |
| Impersonate Empl. – Cargo Hold | 0.7 |
| Impersonate Contr. – Cargo Hold | 0.6 |
| Impersonate Contr. – Galley | 0.6 |
| Landing Gear – Exit | 0.4 |
| Cargo Hold – Exit | 0.2 |
| Galley – Exit | 0.3 |

The optimal path for an intruder (i.e., the path that maximizes the probability of success) is the path of greatest vulnerability to the system. In our simple example, we would compute a probability of successful attack of 0.0034 for an intruder whose strategy is to gain entry to the apron area through the service vehicle gate, then impersonate a contractor (probably a catering service worker) to access the aircraft galley and place the device there before exiting.

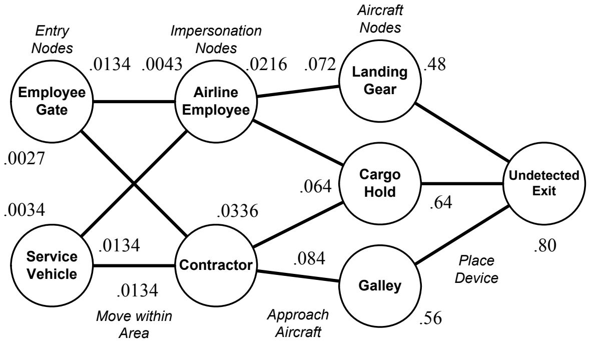

**Figure 5. Summary of intruder strategy and probability of success.**

The existence of this strategy does not mean that all intruders will always proceed in exactly the way indicated. It does mean that if an intruder were perfectly informed, this would be a strategy through which the probability of a successful attack could be maximized. In actuality, the probability of successful attack is likely to be less than this maximum value because intruders will have less-than-complete information and may not optimize their strategy. The solution to the MDP also provides useful information on the conditional probability of success for an attacker that reaches a certain point in the network, regardless of whether or not he/she followed the optimal strategy. For example, if an intruder succeeds in reaching the cargo hold of the aircraft (despite the fact that this is not an optimal strategy), the probability of a successful attack from that point on is 0.064.

# 5. Representing Imperfect Information

One useful representation of imperfect information is to assume that a potential intruder does not know the values of the probabilities in Tables 1 and 2, but has perceptions of those probabilities that contain errors. An intruder with imperfect information will attempt to construct an optimal strategy, but because of errors in perception of detection probabilities, the strategy is likely to actually be suboptimal against the real probabilities. Simulation is an effective tool to explore the effects of imperfect information represented in this way.

Suppose that the perception of a given detection probability is represented as a beta random variable with parameters $a > 0$ and $b > 0$. The mean of such a random variable is $\frac{a}{a+b}$, and the variance is $\frac{ab}{(a+b+1)(a+b)^2}$. If the intruder's perception of an unknown probability $\pi$ is unbiased, $\frac{a}{a+b} = \pi$ , and we can express one of the parameters in terms of the other – e.g., $b = \frac{(1-\pi)a}{\pi}$ . By varying $a$, we can change the variance (i.e., the level of uncertainty in the perception of $\pi$) and set $b$ in terms of $a$ to maintain the same expected value. A convenient way to create experiments is to set the coefficient of variation for the distribution and then solve for the values of $a$ and $b$ that will maintain the desired mean and achieve the required standard deviation. The coefficient of variation for the beta distribution is $\sqrt{\frac{b}{a(a+b+1)}}$ .

Alternatively, we can assume that the intruder's perception of the unknown probability may be biased. If we specify both the coefficient of variation in the distribution and the degree of bias ($\pi - \frac{a}{a+b}$ ), we can solve for values of $a$ and $b$ to satisfy those requirements.

For any setting of the values for the parameters $a$ and $b$, we can sample from the perception distribution to simulate an intruder operating with some specified level of imperfect information. Of course, this concept extends to imperfect information with respect to any number of probability estimates. Replicating this simulated sampling leads to varying choices of paths through the network by the imperfectly informed intruder, each of which has a different probability of success. This allows construction of an estimated probability distribution for the likelihood of successful attack by an intruder operating at that level of imperfect information, as well as a probability distribution over possible paths through the network. The distribution of path choices allows us to reach some conclusions regarding the likelihood that an intruder will appear at certain points in the network.

# 6. Illustrative Simulation Results

To illustrate these ideas, we will consider a series of experiments using the basic network from Figure 1, and compare the results to the perfect-information solution in Figure 5. As a first experiment we assume that the intruder's perception of the detection probabilities (at the nodes and along the arcs) is unbiased, but has a coefficient of variation of 0.1 for all non-zero probabilities (i.e., excluding the first four entries in Table 2).

As an example of the beta distribution parameter computations, consider the detection probability for the arc connecting "Impersonate Contractor" to "Cargo Hold" (the seventh row of Table 2). The true value for this probability is 0.6. To determine the *a* and *b* parameters of the beta distribution to represent imperfect information, we establish the two equations:

$$\frac{a}{a+b} = 0.6 \tag{17}$$

$$\sqrt{\frac{b}{a(a+b+1)}} = 0.1 \tag{18}$$

We then solve for *a* and *b*, leading to the values *a* = 39.4 and *b* = 26.27. This computation is repeated (for different underlying probabilities in equation 17) to produce *a* and *b* parameters for all the non-zero detection probabilities.

In each simulation experiment, the success probability for a given node or arc is adjusted to accommodate the sampled value of the detection probability. The retreat probabilities at the nodes are unchanged. This adjustment ensures that the required probabilities sum to 1.0.

Table 3 summarizes the results of 30 replications of the simulation. The path descriptors use the node numbering scheme from Table 1, and are listed in order of decreasing probability of success. The probabilities of use are rounded to two decimal places, and may not add exactly to 1.0. The path found in the perfect-information case (2-4-7-8) is one of the two most likely paths when the intruder has imperfect information, but approximately 63% of the time, the imperfectly informed intruder will choose a suboptimal path, even when the variability in the perceptions of detection probabilities (as measured by the coefficient of variation) is relatively small (0.1). The average probability of success for an intruder with this level of information is .00279, approximately 17% lower than for the perfect information case. This experiment indicates that even a little reduction in information about the system can have a significant effect on reducing the likelihood of a successful attack.

**Table 3: Summary of results when probability estimates are unbiased and coefficient of variation is 0.1.**

| Chosen Path | Probability of Use | Probability of Success |
|---|---|---|
| 2-4-7-8 | 0.37 | .00336 |
| 1-4-7-8 | 0.17 | .00269 |
| 2-4-6-8 | 0.37 | .00256 |
| 1-4-6-8 | 0.07 | .00205 |
| 2-3-5-8 | 0.03 | .00108 |

In addition to information on average probability of success, the path data and probabilities in Table 3 can be used to estimate the likelihood that an intruder will appear at a given point in the network, given the level of imperfect information hypothesized. This is done simply by summing probabilities for paths that include a given node or arc. For example, we might be particularly interested in the relative likelihoods of attempts to place explosives in the three different areas of the aircraft. In this case, we could use the results in Table 3 to conclude that the probabilities of an intruder attempting to use the landing gear (node 5), the cargo hold (node 6) and the galley (node 7) are .03, .44 and .54, respectively (again rounded to two decimal places).

Further insight into the effects of imperfect information can be obtained by increasing the level of uncertainty. A second experiment increased the coefficient of variation in the detection probability perceptions to 0.3. The perceptions are still considered to be unbiased. Table 4 summarizes the results, again based on 30 replications of the simulation.

Comparing Table 4 to Table 3, we see that the increase in uncertainty about the correct detection probabilities causes the optimal path to be chosen less frequently, and a very suboptimal path (1-3-5-8) appears in the list of possibilities. Overall, the average probability of success is .00266. This is a decrease from the case where the coefficient of variation is 0.1, but only about 5%. In this sample problem at least, a small amount of uncertainty in the perceived detection probabilities is important, but making that uncertainty much larger has relatively little effect on the expected probability of successful attack, as long as the perceptions are unbiased.

**Table 4: Summary of results when probability estimates are unbiased and coefficient of variation is 0.3.**

| Chosen Path | Probability of Use | Probability of Success |
|---|---|---|
| 2-4-7-8 | 0.33 | .00336 |
| 1-4-7-8 | 0.23 | .00269 |
| 2-4-6-8 | 0.27 | .00256 |
| 1-4-6-8 | 0.07 | .00205 |
| 2-3-5-8 | 0.03 | .00108 |
| 1-3-5-8 | 0.07 | .00086 |

There is a somewhat more noticeable effect of the increase in uncertainty on the probabilities of the intruder attempting to use different parts of the aircraft. From the results in Table 4, we can compute estimates of the probability that the intruder would attempt to use the landing gear (node 5), the cargo hold (node 6) and the galley (node 7) as 0.1, 0.34, and 0.56, respectively. There is a noticeable shift in likelihood from the cargo hold to the landing gear for less well-informed intruders. This insight can be helpful to security forces.

To test the effects of biased perceptions, we have conducted a third simulation experiment. The coefficients of variation in the detection probability perceptions are set to 0.1, as in the first experiment, but we introduce a bias on two of the perceived probabilities – the detection probabilities associated with a contractor approaching the aircraft, either the cargo hold or the galley. In Table 2, the "true" values are indicated to be 0.6, but we assume that the intruder believes (on average) that the values are 0.9 for both probabilities. Intuitively, we expect that these misperceptions will tend to drive the intruder's attack path away from paths that use those two arcs, and since one of the two arcs is part of the optimal path under perfect information, the net effect should be a reduction in success probability for the intruder.

Table 5 summarizes the results of the experiment, again based on 30 simulation replications. The overall average probability of success for an attack is reduced to .00144, a reduction of 48% from the value in experiment 1 (.00279), and a reduction of 57% from the original value based on perfect information. The misperception of detection probabilities on the two arcs makes it much less likely that the intruder will attempt to use those arcs (probability of 0.23 versus 0.97 in the first experiment). Attacks are much more likely to be focused on paths (and areas of the aircraft) where the real detection probability is higher, leading to much lower success probability for the intruder. In the results shown in Table 5, the probability of the intruder attempting to use the landing gear area is 0.37, as compared to 0.03 in experiment 1, and the probability of attempts through the galley has decreased from 0.54 to 0.1.

The level of bias in the perceptions of the detection probabilities on arcs 4-6 and 4-7 used in this experiment is substantial, and smaller assumed biases would create less dramatic results. However, we have only introduced the bias on two arcs in the network. More widespread misperceptions would be likely in a larger system. This experiment does indicate that creating biased perceptions of detection probabilities among potential intruders can be very effective in reducing the likelihood of successful attacks by "steering" those attacks into areas where detection really is very likely.

**Table 5: Summary of results when probability estimates are biased on arcs 4-6 and 4-7.**

| Chosen Path | Probability of Use | Probability of Success |
|---|---|---|
| 2-4-7-8 | 0.10 | .00336 |
| 2-4-6-8 | 0.13 | .00256 |
| 2-3-5-8 | 0.33 | .00108 |
| 2-3-6-8 | 0.33 | .00096 |
| 1-3-5-8 | 0.03 | .00086 |
| 1-3-6-8 | 0.07 | .00077 |

There are several means through which a system operator might create such misperceptions. Implementing inexpensive, highly visible (though perhaps not really very effective) detection mechanisms might be one means. Supplying disinformation about real operations or procedures may be another, although this has obvious drawbacks as well.

# 7. Optimizing Resource Allocation for Security Improvement

The illustrative analysis leads us to an obvious question: If it were possible to estimate a cost function for changes within the network that would reduce the likelihood of a successful intrusion, could we identify the most effective (i.e., minimum cost) way of achieving a desired (small) probability of successful intrusion? This question can be answered using a bi-level optimization formulation. At the "upper level" we have an optimization that determines changes at nodes and arcs in the network so as to minimize cost, subject to a constraint that the resulting probability of successful attack is no greater than a specified value. However, the probability of successful attack is determined as the solution to a "lower level" optimization (optimizing the intruder's strategy, given the characteristics of the network he/she is facing and his/her level of knowledge about that system).

To be more specific about this optimization, consider again the model of the intruder's strategy expressed in equations (11)-(13). There are at least five ways that the system operator (or "defender") can act to reduce the likelihood that the intruder will be successful:

- Increase the probability of detection at barrier (node) $i$ ; this might be accomplished either by increasing the sensitivity of the detection process, or by increasing the time required to penetrate the barrier, allowing the existing detection mechanisms more time to be effective.
- Increase the probability of detection on movement arcs *ij* between nodes.
- Add new barriers that must be negotiated; this is represented by a new node in the network, with reconnection of existing arcs to force some (or all) intruders' paths to go through the new node.
- Remove existing arcs in the network; this represents some additional constraints (either physical or virtual) on movement within the system.
- Reduce the level of information that potential intruders have about the system structure and detection probabilities, creating additional uncertainty for the intruders, and perhaps some level of "disinformation" that would lead them to make poor choices in their attack strategy.

From the standpoint of the model we have defined, the third and fourth strategies listed can be considered to be special (extreme) cases of the first two strategies. The fifth strategy is quite different from the first two, and is analyzed via the methods described in Sections 5 and 6.

Consider, for the moment, the first two strategies for reducing the vulnerability of the system (implicitly including the third and fourth as well). Suppose that the initial detection probability at node $i$ is denoted $d_i^0$ , and the increase in that probability is denoted $\Delta_i$ , so that the actual detection probability in effect is $d_i = d_i^0 + \Delta_i$ .

Similarly, we will assume that the initial detection probability on arc *ij* is $\delta_{ij}^0$ , and the increase in that probability is $\gamma_{ij}$ , so the actual detection probability in effect is $\delta_{ij} = \delta_{ij}^0 + \gamma_{ij}$ . Increases in the detection probabilities are assumed to require expenditures $C_i(\Delta_i)$ and $K_{ij}(\gamma_{ij})$ . In the current formulation, the cost functions are separable by node and arc, but a more general cost function could be used without changing the structure of the bi-level optimization formulation.

We will use $E$ to denote the set of entry nodes to the system network, and then express the "upper level" problem as follows:

$$\text{Min } \sum_i C_i(\Delta_i) + \sum_{ij} K_{ij}(\gamma_{ij}) \tag{19}$$

subject to:

$$w^*(i) \le W^* \qquad \forall\, i \in E \tag{20}$$

$$d_i = d_i^0 + \Delta_i \qquad \forall\, i \tag{21}$$

$$\delta_{ij} = \delta_{ij}^0 + \gamma_{ij} \qquad \forall\, ij \tag{22}$$

$$\Delta_i \ge 0 \qquad \forall\, i \tag{23}$$

$$\gamma_{ij} \ge 0 \qquad \forall\, ij \tag{24}$$

In (20), the $w^*(i)$ values are the optimal solution to the "lower level" problem, specified as follows:

$$\min \sum_i \beta_i w(i) \tag{25}$$

subject to:

$$w(i) - \sum_j P_{ij}(a_i \mid d_i, \delta_{ij}) w(j) \ge 0 \qquad \forall\, i \ne g,\, a_i \tag{26}$$

$$w(g) - \sum_j P_{gj}(a_g \mid d_g, \delta_{gj}) w(j) \ge 1 \qquad \forall\ \ a_g \tag{27}$$

$$w(i) \ge 0 \qquad \forall\, i \tag{28}$$

In (26) and (27), the transition matrix is written as $P_{ij}(a_i \mid d_i, \delta_{ij})$ to reflect the fact that it depends on the values of $d_i$ and $\delta_{ij}$ determined in the upper problem. The lower problem in (25)-(28) is the same problem as in (11)-(13), but is re-written to reflect the specific knowledge of $R_i(a_i)$ values that relevant to this problem, and to emphasize its connection to the upper problem in (19)-(24).

A solution procedure for this bi-level optimization searches over possible values of $\Delta_i$ and $\gamma_{ij}$, and for each set of values, solves the lower problem to find $w^*(i)$ (after translating the $d_i$ and $\delta_{ij}$ values into a new transition matrix $P_{ij}(a_i \mid d_i, \delta_{ij})$). A general issue (which is endemic to bi-level models) is that it is difficult to guarantee convergence of solution algorithms to true optimal solutions in the upper model. Bard [1] describes this general difficulty.

# 8. Extensions

Several possible extensions to this analysis are possible. First, the bi-level optimization problem formulated in Section 7 is likely to be difficult to solve to optimality, and investigation of potential solution methods is an important direction for further work.

Second, other aspects of imperfect intruder information could be included, such as imperfect knowledge about what barriers (nodes) and arcs exist in the system. This type of imperfect information can be incorporated into the general analysis framework described in this paper.

A third useful extension is to consider where improvements in security (i.e., increases in detection probability) would be most effective against several classes of potential intruders (i.e., intruders with differing levels of information about the system).

A fourth useful extension is to create semi-Markov models for the processes of attempted penetration of barriers. This would allow more accurate representation of the uncertain time required to penetrate a given barrier, as well as offer the opportunity for time-dependent detection probabilities (i.e., the longer an intruder is present at a barrier, the more likely it becomes that he/she will be detected). This extension could improve the range of applicability of the model.

# 9. Conclusions

The objective of the analysis presented here is to provide guidance to system owners and operators regarding effective ways to reduce vulnerabilities of specific infrastructure facilities. To accomplish this, we have developed a Markov Decision Process (MDP) model of how an intruder might try to penetrate the various barriers designed to protect the facility. The solution to this MDP model provides insight into the level of vulnerability of the facility (the probability of successful intrusion) and indicates where the vulnerabilities are (the most likely paths for the intruder).

Lower level models of intruder detection at barriers (nodes) of the system can be built as Hidden Markov Models, and the results of those lower level models can be aggregated for use in the MDP of intruder strategy for attacking the system. A key aspect of this analysis is representing imperfect information on the part of the intruders. Simulation is used as a tool to evaluate the effects of varying levels of imperfect information, sampling from distributions of detection probabilities and using those samples to construct distributions of intruder path choices through the network and overall success probability.

A small example problem illustrates that even relatively small amounts of uncertainty in the information the intruders have about the system can significantly affect the probability that they can mount a successful attack. If the uncertainty is combined with bias in the perceptions of some system parameters, the effect on the intruders is magnified. In the small example studied, biased perceptions of two key detection probabilities combined with small amounts of uncertainty in perceptions of all the detection probabilities reduces the likelihood of a successful attack by a factor of about two. In addition to allowing us to estimate the probability of a successful intrusion, the simulation also allows us to estimate the likelihood of attacks appearing at specific locations in the network. This is very useful information for security forces.

The intruder model also provides the basis for consideration of possible strategies to reduce the probability of a successful attack on the facility. The process of searching for cost-effective strategies to reduce system vulnerability can be formally cast as a bi-level optimization problem, as discussed in section 7. This provides a promising direction for further work.

Successful implementation of the model described in this paper depends very directly on two important tasks: 1) constructing large-scale networks that represent the various barriers and movement possibilities in a system; and 2) estimating the various probabilities embedded in the *A* and *B* matrices that are elements of the HMM's at each network node. Quite clearly, if the constructed network does not reflect accurately the barriers to intrusion and possible paths for intruders, the resulting computations from the model will be flawed. Constructing an accurate network representation requires significant system knowledge and also the ability to "think like an attacker." Estimating the probabilities is also a challenging task. There are tools that have been created for estimating HMM matrices in other application contexts, and the experience gained in those other contexts should provide important insight for this task.

The process of testing, implementing and enhancing the model is an ongoing one, with the expectation that this approach will become an important new tool for the protection of critical infrastructure facilities.

# 10. References

[1] Bard, J.F., “Some Properties of the Bilevel Programming Problem,” *Journal of Optimization Theory and Applications*, 68:2, 1991, 371-378.

[2] Brémaud, P., Markov Chains: Gibbs Fields, Monte Carlo Simulation, and Queues, Springer-Verlag, New York, 1999.

[3] Carlson, R.E., Turnquist, M.A. and Nozick, L.K., *Expected Losses, Insurability and Benefits from Reducing Vulnerability to Attacks*, Report SAND2004-0742, Sandia National Laboratories, Albuquerque, NM, 2004.

[4] Executive Office of the President, *National Strategy for Homeland Security*, July 2002, available on line at http://www.dhs.gov.

[5] Jha, S., Sheyner, O., and Wing, J.M. “Two Formal Analyses of Attack Graphs,” *15th IEEE Computer Security Foundations Workshop*, June 2002, Cape Breton, NS, Canada, 49-63.

[6] Katsikas, S.K., Gritzalis, D., and Spirakis, P., “Attack Modelling in Open Network Environments,” *Communications and Multimedia Security II*, 1996, 268-277.

[7] Katsikas, S.K., Spyrou, T., Gritzalis, D., and Darzentas, J., “Model for Network Behaviour under Viral Attack,” *Computer Communications,* 19:2, 1996, 124-132.

[8] Ourston, D., Matzner, S., Stump, W., and Hopkins, B., “Applications of Hidden Markov Models to Detecting Multi-stage Network Attacks,” *36th Hawaii International Conference on Systems Science*, IEEE Computer Society, Hawaii, 2003, CD-ROM, 10p.

[9] Ourston, D., Matzner, S., Stump, W., and Hopkins, B. “Coordinated Internet Attacks: Responding to Attack Complexity,” *Journal of Computer Security*, 12:2, 2004, 165-190.

[10] Phillips, C.A., and Swiler, L.P., “A Graph-Based System for Network Vulnerability Analysis,” *Proceedings of the 1998 New Security Paradigms Workshop*, Association for Computing Machinery, 1998, 71-81.

[11] President’s Commission on Critical Infrastructure Protection, *Critical Foundations: Protecting America’s Infrastructures*, The White House, Washington, DC, 1997.

[12] Puterman, M.L. *Markov Decision Processes*. Wiley, New York, 1994.

[13] Ross, S.M., *Introduction to Probability Models*, 7th Edition, Academic Press, San Diego, CA, 2000.

[14] Sheyner, O., Haines, J., Jha, S., Lippmann, R., and Wing, J.M., “Automated Generation and Analysis of Attack Graphs,” *Proceedings of the IEEE Computer Society Symposium on Research in Security and Privacy*, Berkeley, CA, May 2002, 273-284.

[15] Soh, B.C., and Dillon, T.S. “Setting Optimal Intrusion-Detection Thresholds,” *Computers & Security*, 14:7, 1995, 621-631.

[16] Swiler, L.P., Phillips, C.A., Ellis, D., and Chakerian, S., “Computer Attack Graph Generation Tool,” *Proceedings of the 2nd DARPA Information Survivability Conference and Exposition*, 2001, 307-321.

[17] Warrender, C., Forrest, S. and Pearlmutter, B. “Detecting Intrusions Using System Calls: Alternative Data Models,” *Proceedings of the 1999 IEEE Symposium on Security and Privacy*, 1999, 133-145.

# Distribution:

| | | | |
|---|---|---|---|
| 5 | MS1138 | Dean A. Jones | 6223 |
| 1 | MS1138 | Chad E. Davis | 6223 |
| 1 | MS1138 | Karen S. Rogers | 6223 |
| 1 | MS1138 | John L. Mitchiner | 6220 |
| 2 | MS9018 | Central Technical Files | 8944 |
| 2 | MS0899 | Technical Library | 4536 |
| 1 | MS0188 | D. Chavez, LDRD Office | 1030 |