SAND2009-3568C

Risk-Based Cost-Benefit Analysis for Security Assessment Problems

Gregory D. Wyss, John Darby, Consuelo Silva, and Andrew Walter
Sandia National Laboratories – Security Systems Analysis Department
PO Box 5800, MS 0757, Albuquerque, NM 87185-0757
gdwyss@sandia.gov (505) 844-5893

ABSTRACT

Decision-makers cite the need to perform risk-based cost-benefit analyses to prioritize security investments. But the most common performance metric for physical security systems is poorly suited to cost-benefit analysis because comparable changes in adversary characteristics can produce dramatically different changes in the metric and lead the decision-maker toward biased or questionable investment decisions. This paper describes ongoing work to define a new physical security effectiveness metric based on the resources required for an adversary to be successful when executing his or her most advantageous attack scenario. This metric is compatible with traditional cost-benefit optimization algorithms, and can enable the development of an objective risk-based cost-benefit method that will enable security investment option prioritization. It also enables decision-makers to more effectively communicate the justification for their investment decisions with stakeholders and funding authorities.

INTRODUCTION

For many years, safety investment decisions have been made using risk-based cost-benefit analysis in which the benefit metric is heavily based on a quantitative estimate of risk reduction. Many seek to perform similar analyses to prioritize security investments. However, for high-security facilities, security risk is much harder to quantify than safety risk since the probability of attack is highly uncertain and depends strongly on unquantifiable psychological factors such as deterrence and adversary goal intensity. In addition, the most common performance metric for physical security systems is "probability of effectiveness at the design basis threat" (P_E at DBT), which represents the probability that a design basis adversary will fail to achieve his objectives given that he initiates his most advantageous attack scenario. This metric is poorly suited to cost-benefit analysis because seemingly small changes in adversary characteristics can profoundly affect P_E when the threat is near a system's breaking point. Also, systems with many attack pathways exploitable by less-than-DBT-level adversaries will continue to have low overall P_E even as investments dramatically raise the difficulty of an attack. These characteristics make it difficult to prioritize security investment options on the basis of P_E – especially across multiple targets or facilities. For these reasons, risk-based cost-benefit analysis is often very subjective for security investment decisions, which can lead to biased or questionable security investment decisions.[*, 1, 2]

[*] The need for a new process has been noted by both external review panels and NNSA decision-makers. For example, in 2002, the Commission on Science and Security recommended that DOE "Develop and practice risk-based security." The commission specifically proposed "the establishment of a risk-based systems approach to the development, analysis, and implementation of security policies throughout the DOE complex." [1] In addition, at the June 2004 NNSA Security Summit, Admiral (ret.) Mies stated that "An enterprise approach to security is missing." [2]

In order to overcome these obstacles and enable risk-based cost-benefit security investment prioritization, Sandia National Laboratories has begun a Laboratory Directed Research and Development project to develop a risk-based cost-benefit analysis method for security investment prioritization that focuses on the adversary's perspective. The project team views adversary attack preparation activities as a project planning exercise, wherein a planner has success criteria (e.g., cause specific consequences) and chooses among alternative courses of action that meet these criteria (plausible attack scenarios) while considering the resources required to ensure a high likelihood of success. An investment reduces security risk to the degree that it increases the resources required for an adversary to be successful when executing his most advantageous attack scenario (which may be a different scenario after the investment is completed) By quantifying this increased degree of difficulty as a measure of risk, an objective risk-based cost-benefit method can be developed that would enable security investment option prioritization using traditional cost-benefit optimization algorithms. Such a method would enable decision-makers to achieve better balance among competing security interests (e.g., multiple facilities), provide more objective and unbiased justification for investment decisions, and reduce second guessing of investment decisions by funding authorities.

A concrete but notional example of the issues involved in security cost-benefit can be observed in the graphs in Figures 1 and 2. Figure 1 shows the degree of difficulty associated with an adversary's most advantageous attack at five hypothetical sites, expressed as a comparison between the design basis threat (DBT) and the resources required for an adversary to be successful in his most advantageous attack. Figure 2 indicates the estimated P_E for those same scenarios against a DBT adversary force. Figure 1 indicates that the security at Site B significantly exceeds the requirements of the DBT, Site C has relatively minor security deficiencies, but Site D would be dramatically easier for an adversary to attack than any other site. Assuming comparable consequences, Site D has the most pressing security improvement needs, and one might consider reducing security costs at Site B. However, examining P_E for each site gives a very different impression: Sites A, B, and E all look comparably good, and Sites C and D appear to have comparable security improvement needs. Thus, one can see the improvement in overall security as mitigation measures make attacks more difficult at the most vulnerable sites, but using P_E as the primary metric in a security cost-benefit analysis can lead a decision-maker to make inappropriate investment decisions.

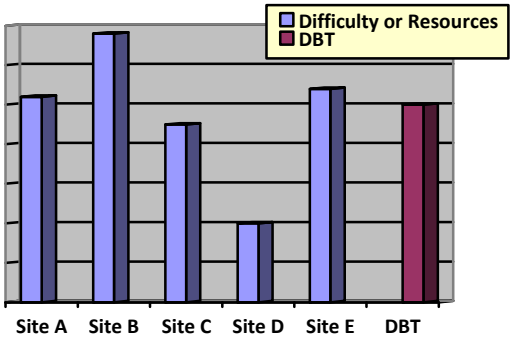

Figure 1. Degree of difficulty for an adversary's most advantageous attack at five hypothetical sites, compared with the design basis threat.

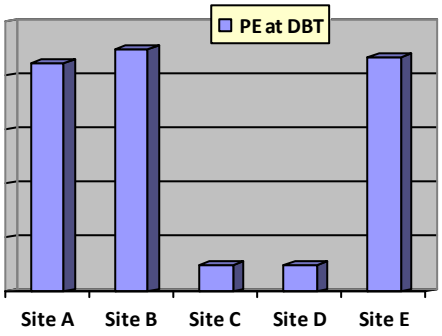

Figure 2. Estimated P_E for the hypothetical scenarios identified in Figure 1 against a design basis threat adversary force.

Because this project is still in progress, this paper describes recent work in the development of this security risk-based cost-benefit analysis method, including a comparison to traditional risk assessment methods and potential metrics to represent required adversary resources.

DEFINITIONS OF RISK

If risk is to be the basis for cost-benefit decision-making, then the definition of risk must be compatible with the decision-making process. For this reason, we must examine the common definitions of risk and adjust them as necessary to better support the security decision-making process.

Current Definitions of Risk

The most commonly cited definition of risk is loss expectancy, in which one multiplies the likelihood of a scenario (its probability or frequency) by the expected consequences given that the scenario occurs in order to obtain a numerical estimate for risk. Given the probabilistic nature of this computation, one can sum risk over all mutually exclusive and statistically independent scenarios to estimate the "total risk" for a system. This total risk or total loss expectancy is highly relevant for insurance companies and is widely used in safety and security decision-making. This definition of risk carries an implicit assumption that often goes unstated and unrealized: two scenarios may have equivalent risk even though one is a very common, low-consequence event and the other is an exceedingly rare but catastrophic event, as long as they have the same value for the product of likelihood and consequence. This perspective of equivalence is not always shared by stakeholders and decision-makers. It is in part because of this implied equivalence that Kaplan and Garrick, in their seminal article on modern risk assessment methods,[3] asserted that this definition of risk is misleading.

Kaplan and Garrick asserted that risk should not be reduced to a single value – either for a system or even for an individual scenario. Rather than risk being "probability *times* consequence," they assert that risk is "probability *and* consequence." They state, "Fundamentally... a risk analysis consists of an answer to the following three questions:

1. What can happen?
2. How likely is it that [it] will happen? and
3. If it does happen, what are the consequences?

"To answer these questions we would make a list of outcomes or 'scenarios' [where each line in the list] can be thought of as a triplet $<s_i, p_i, c_i>$ where s_i is a scenario identification or description; p_i is the probability of that scenario; and c_i is the consequence or evaluation measure of that scenario, i.e., the measure of damage. If this table contains all the scenarios we can think of, we can then say that it (the table) is the answer to the question and therefore is the risk." Therefore, risk is defined as a collection of such triples, and since each scenario is associated with a probability, one can summarize this set of triples as a "risk curve" which satisfies the definition of a statistical complementary cumulative distribution function.

Difficulties with the Current Definition

Even this modern definition of risk is somewhat problematic when one considers security for high-consequence facilities. Scenarios must be mutually exclusive and statistically independent if one is to aggregate risk according to either of the above risk definitions. In the world of physical security, this condition is clearly not met because an intelligent and malevolent adversary chooses among these scenarios and selects the one that is in his or her best interest. In fact, he even chooses among scenarios that are not represented in our scenario set because he may be deterred from attacking a target at our facility and choose to mount an attack on a different facility altogether. In addition to these mathematical problems, certain practical problems also exist with this definition. Fundamentally, the probability of an attack can only be estimated in a Bayesian sense and is enormously uncertain because we cannot know the true intentions of all adversary groups. Furthermore, this probability can change wildly over time as adversary groups are influenced by local and global political and social events of which we may not even be aware. Therefore, for malevolent human events, a definition of risk that involves the probability of an attack will doom decision-makers to decide on risk mitigation investments under enormous numerical uncertainties that are in fact caused by the very definition of risk.

In order to move past the inevitable and unresolvable arguments regarding the probability of an attack, most security analysts evaluate conditional risk – that is, the risk that exists given that an attack occurs (or, conditional upon the attack occurring). Conditional risk is expressed in terms of the probability that the adversary's attack successfully causes the consequence and the value of the consequence itself. The probability of adversary success is the complement of the widely recognized security performance metric P_E. To assess P_E, one must assume that an attack is carried out by an adversary with particular characteristics (e.g., number of attackers, weapons, tools, etc.), which is often used as a boundary condition for the design and analysis of the security system (a "design basis threat"). After all, for the same scenario, P_E may be very high if the attacker force is two people with shotguns, but very low against a coordinated attack from several Special Forces squads. Security analysts have often observed that seemingly small changes in adversary characteristics can profoundly affect P_E when the threat is near a systems breaking point, and, conversely, seemingly large changes in adversary characteristics can have minimal effect on P_E in other parts of the attack space. In addition, historical attacks indicate that an adversary will assemble resources that the adversary believes are sufficient to ensure a high likelihood of a successful attack, or the adversary will not attempt the attack. Given this adversary behavior, speaking of probabilities is not as useful in the context of an intelligent and malevolent adversary as it is for the random events that comprise safety risk analyses.

Extending the Definition of Risk

In order to overcome the obstacles related to the use of probabilities with malevolent adversaries, we propose a modified definition of risk where, instead of considering the highly uncertain likelihood or probability of a scenario, one considers its difficulty to an adversary. Thus, a security risk analysis consists of answers to the following three revised questions:

1. What can happen?
2. How difficult is it for an adversary to make this event happen? and
3. If an adversary causes this event to happen, what are the consequences?

The triplet for security risk then becomes $<s_i, d_i, c_i>$ where d_i is the degree of difficulty for an adversary to accomplish scenario s_i and cause consequence c_i. This definition explicitly acknowledges the observed adversary attack planning behaviors described above, and addresses the problems associated with using probabilities to describe the intentional actions of both known and unknown intelligent actors. It also supports more robust risk-based cost-benefit decision-making for security applications because, as demonstrated in the Introduction to this paper, the degree of difficulty associated with an adversary's most advantageous attack is a much more useful metric for security decision-making than is P_E or any of the other common probability-based metrics.

While it is easy for an analyst to *describe* the difficulties inherent in a specific attack scenario, these difficulties are hard to express as a single metric – either qualitative or quantitative – because of the large number of disparate factors that may cause difficulty to an attacker. These factors include resources (personnel, expertise, weapons, tools, money, etc.), scenario complexity, knowledge and information about the target and facility, unpredictability of the attack environment, operational security prior to the attack, detectability of pre-attack surveillance, rehearsal, weapon procurement, and so forth. Therefore, a metric or system of metrics must be developed to describe and summarize the degree of difficulty for an attack scenario if one is to successfully apply this proposed definition of risk. However, with such a metric, this definition of risk can form the basis of an objective risk-based cost-benefit analysis method that would enable security investment prioritization using traditional cost-benefit optimization algorithms.

RANKING SCENARIOS ACCORDING TO DIFFICULTY

To date, our research has not uncovered any generally accepted metric or system of metrics to answer the question, "How difficult is it for an adversary to accomplish this scenario?" or the related question, "How much more capable is one adversary compared to another?" However, a few methods have been proposed for other purposes that may be applicable to this problem.

Simple Scenario Ranking Systems

The most straightforward method for ranking scenarios is a qualitative or semiquantitative comparison with capabilities of predefined adversaries (either actual or notional). Garcia[4] encourages security analysts to collect and organize threat information for various types of adversaries (e.g., terrorist, criminal, extremist) when designing and evaluating physical protection systems. As the capabilities of these adversaries are compiled, it is possible for experts to rank the adversaries from least capable to most capable, and then to evaluate and rank scenarios on the basis of which adversary has the threshold capabilities to be able to successfully execute the scenario in question. A similar ranking system can be devised when security systems are evaluated against multiple threat levels. For example, one might devise a low-, medium-, and high-level threat for evaluation purposes during a facility security analysis, or one might establish different "design basis" threat levels for the protection of different types of targets, as DOE has done in its Graded Security Protection Policy.[5] Scenarios can be ranked on the basis of which notional threat would find the scenario favorable for planning and execution in view of their tactical or training advantages and resource usage. A notional scale is presented in Table 1.

Table 1. Notional scale for ranking scenarios against multiple hypothesized adversaries or threats.

Scenario Difficulty Increases as Adversary Thresholds are Reached ➔	
	A high-level adversary would be unlikely to be successful in accomplishing this scenario
	A high-level adversary would find this scenario acceptable but not necessarily desirable
	A medium-level adversary would be unlikely to be successful in accomplishing this scenario, but a high-level adversary would find this scenario favorable
	A medium-level adversary would find this scenario acceptable but not necessarily desirable
	A low-level adversary would be unlikely to be successful in accomplishing this scenario, but a medium-level adversary would find this scenario favorable
	A low-level adversary would find this scenario acceptable but not necessarily desirable
	A low-level adversary would find this scenario favorable

A similar adversary ranking system and scenario categorization system has been developed for the evaluation of cyber security. In this work, a series of seven[6] or eight[7] threat levels have been characterized and ranked by experts based on the adversary capabilities in a number of dimensions that are relevant to cyber attacks, such as stealth, time that an adversary will dedicate to planning and executing an attack, level of access to the target, number of personnel dedicated to planning and executing the attack, and so forth. In Reference 7, the authors compare the characteristics of several observed attacks with the characteristics seen in the developed adversary threat-level matrix in order to validate the applicability of the generic threat matrix. This activity could also be carried out on hypothesized attacks during a system design and evaluation process, which has strong parallels to the risk definition and cost-benefit analysis method proposed in this paper.

The DOE Adversary Mission Analysis Tool

The DOE Vulnerability Assessment Technical Standard,[8] Chapter VII, presents a tool called Adversary Mission Analysis as an expert-based approach to help analysts sift through the large number of scenarios that can be generated in a vulnerability analysis and prioritize which scenarios should be subjected to more detailed and expensive neutralization modeling. This methodology was derived from the process described in the U.S. Army's FM101-5, *The Military Decision-Making Process*. The tool evaluates scenarios according to several dimensions, such as

- the type of critical intelligence required for adversary mission success and the degree of difficulty inherent in collecting that intelligence,
- the likelihood of the adversary mission being compromised prior to the beginning of the attack,
- the simplicity and flexibility embodied in the attack scenario,
- reliance of the attack scenario on advanced skills, special equipment, or rare assets, and
- the favorability of the attack scenario compared with the features of the facility security system.

The tool provides a scoring system based on a weighted sum of scores in each of nine dimensions, with higher scores being more advantageous to the adversary and representing, in the view of the authors, the most likely adversary courses of action. The highest scoring scenarios "are to be used

to develop the 'representative-case' scenarios [as they] have the highest chance of being successfully executed." Scenarios scoring in the mid-range "require some level of sensitivity analysis to ensure no hidden impacts arise." If the scoring can be carried out consistently across systems and across sites, possibly through the use of a peer review process, it may be possible to use these scores as surrogates for the degree of scenario difficulty in a security cost-benefit analysis method.

Other Scenario Scoring Options

A more detailed but complex scenario scoring system has been proposed by Walter.[9] In this work he proposes that one could assign a point value to each tool, weapon, expertise, and attacking individual in order to compare the adversary resources used in various observed attacks (both successful and unsuccessful) with those that have been brought to bear against DOE sites during various security design and evaluation exercises. As currently envisioned, this "Points-Based Design Basis Threat" would be a way to ensure that security design and evaluation exercises are focused on "reasonable" adversaries instead of "extreme" adversaries that are "10 feet tall." However, one could use this type of scoring system in conjunction with the risk definition and security cost-benefit analysis method described in this paper if one were to evaluate scenarios on the basis of the minimum set of adversary resources required for an adversary to be successful when executing the attack scenario. One could then score this minimum set of required resources using this type of point system as the basis for determining the degree of difficulty for the scenario.

The project team is continuing to examine existing methods for ranking scenarios according to their degree of difficulty. The objective is to develop an initial workable difficulty metric that will support the enhanced definition of risk for use in demonstrating a risk-based cost-benefit decision-making method. Given the intricacies and subtleties of the disparate dimensions involved in this ranking process, the authors believe that the difficulty metric will need to be revised and refined well beyond the completion of this research project.

FUTURE WORK

To date, work on this project has focused on the two areas described above: development of an enhanced definition for risk that is applicable to malevolent human activities, and implementation of that definition by developing a system of metrics that can be used to represent the degree of difficulty for a scenario. The next step of the project will entail demonstration of this risk metric in a risk-based cost-benefit analysis method. With this revised definition of risk, the team believes that the cost-benefit analysis method development will be straightforward. The revised risk definition and metric would be simply a new measure for benefit that can be incorporated within existing cost-benefit analysis methodologies with little or no adaptation. In addition, the development of a cost-benefit optimization method is also believed to be straightforward because the basic framework described in Figure 1 can be readily used in conjunction with existing and well-studied cost-benefit optimization methods and tools.

A separate task is being pursued under the project to automate the search for the minimum resources required for the adversary to succeed. The project team is examining the potential to implement any ranking method for scenario difficulty or adversary capabilities as part of the utility function that is evaluated within an optimization routine. Conceptually, the optimization routine can interrogate a security assessment model such as DANTE in order to enable automated assessment of the threshold capabilities and resources required for an adversary to successfully

complete an attack scenario. This would help to address criticisms that the scenario ranking methods are primarily based on expert judgment and, thus, subject to the bias of human analysts.

The risk-based cost-benefit analysis method described above applies principally when the scenarios under consideration all lead to comparable consequences. The situation becomes much more complicated when widely varying consequences are considered, such as in the case of a military base commander who is charged with protecting a base commissary, base housing, and base weaponry. For such a situation, numerical optimization may be very difficult, but the philosophy embodied in the enhanced risk definition can be used to provide guidance for the most appropriate security investments. Consider Figure 3. In this graph, each scenario is represented by a dot according to the enhanced risk definition described above—that is, its degree of difficulty (as represented by the resources required for the adversary to succeed) and its expected consequences. The line that forms the upper left boundary of this set of points represents the most optimal adversary attack scenarios. While one cannot predict which scenario an adversary might select, a rational adversary will not select a scenario that is a great distance from this line, and if an adversary does select a scenario that is far from this line, the facility owner should be privately thankful that the adversary did not choose to apply these same resources toward a scenario that would result in significantly greater consequences.

A risk management methodology that is based on this concept is illustrated in Figure 4. Notionally, the objective is to move the optimal attack line (the "risk frontier") down and to the right over time. This makes the adversary's best attack options either harder to accomplish (by reducing vulnerabilities, which moves scenario dots to the right), less consequential (by reducing the magnitude and/or likelihood of expected consequences given a successful attack, in effect moving scenario dots down), or both. In this figure, risk managers select projects with the intent to remove all scenario dots from the red area within 5 years and from the yellow area within 10 years. The project team intends to examine this representation of risk and its application to security risk management during the remainder of the project.

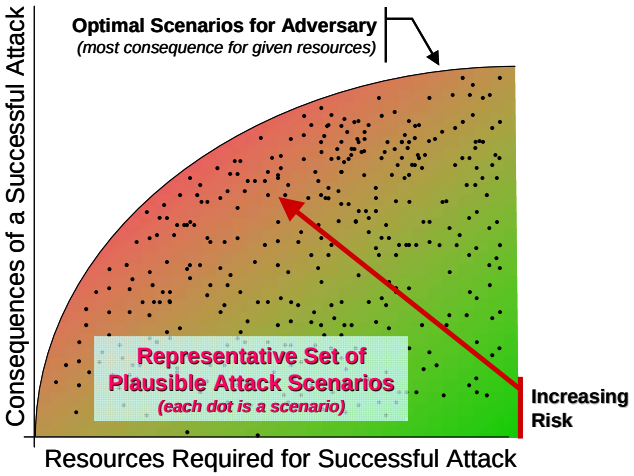

Figure 3. Security risk, viewed as a function of consequences and adversary requires required for a successful attack.

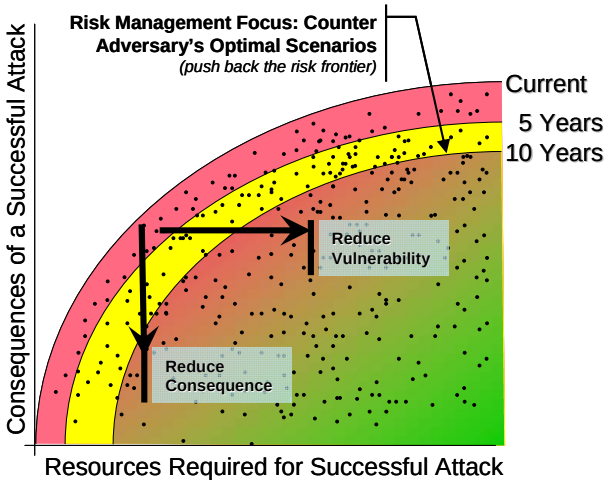

Figure 4. Notional future risk management strategy.

Summary and Conclusions

This paper has described ongoing work in a Laboratory Directed Research and Development project to develop a risk-based cost-benefit analysis method for security investment prioritization. The project team found that the most common security metric, P_E at the design basis threat, is not suitable for use in cost-benefit analysis because it can mislead a decision-maker to making inappropriate investment decisions. In order to enable risk-based decisions, the project team developed an enhanced definition for risk that considers the degree of difficulty an adversary would experience in attempting to accomplish a scenario instead of the traditional questions about attack likelihood or success probabilities. The degree of difficulty for an attack scenario is a useful and robust metric for security cost-benefit analysis. While there is not a generally accepted metric or system of metrics to rank scenarios according to their degree of difficulty or to compare the capabilities of one adversary with another, several methods have been proposed to develop such metrics. The project expects to develop an initial metric in order to demonstrate the usefulness of the risk-based cost-benefit analysis method for investment prioritization decisions and to demonstrate its compatibility with existing cost-benefit optimization tools. Thus, we conclude that risk-based cost-benefit analysis is possible and will represent a significant advantage for decision-makers as they seek to optimize the benefits of their security investments.

ACKNOWLEDGMENTS

This work is being performed using funding from the Laboratory Directed Research and Development Program at Sandia National Laboratories. The authors are grateful for this support. 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. This paper is approved for unlimited release as SAND2009-XXXX

REFERENCES

[1] Recommendation 3, *Science and Security in the 21st Century*, Commission on Science and Security (John Hamre, Chairman), April 2002.

[2] Admiral (ret.) Mies' comment regarding his ongoing review of NNSA security, NNSA Security Summit, June 2, 2004.

[3] Kaplan, S., and Garrick, B.J., "On the Quantitative Definition of Risk," *Risk Analysis*, **1**:1(11), 1981.

[4] Garcia, M.L., *The Design and Evaluation of Physical Protection Systems*, Butterworth-Heinemann (Elsevier), Burlington, MA, 2001.

[5] US Department of Energy, Graded Security Protection Policy (U), DOE O 470.3B, US Department of Energy, Washington, DC, August 12, 2008.

[6] Duggan, D.P., *Generic Threat Profiles*, SAND2005-5411, Sandia National Laboratories, Albuquerque, NM, July 2005.

[7] Duggan, D.P., Thomas, S.R., Veitch, C.K.K., and Woodard, L., *Categorizing Threat: Building and Using a Generic Threat Matrix*, SAND2007-5791, Sandia National Laboratories, Albuquerque, NM, September 2007.

[8] US Department of Energy, *DOE Standard: Vulnerability Assessment*, Chapter VII, "Adversary Mission Analysis," DOE-STD-0005-2008 (Draft), Official Use Only, US Department of Energy, Washington, DC, September 2008.

[9] Walter, A., *A Resource Point-Based DBT/ACL*, Official Use Only white paper, Security Systems Analysis Department, Sandia National Laboratories, Albuquerque, NM, September 10, 2007.