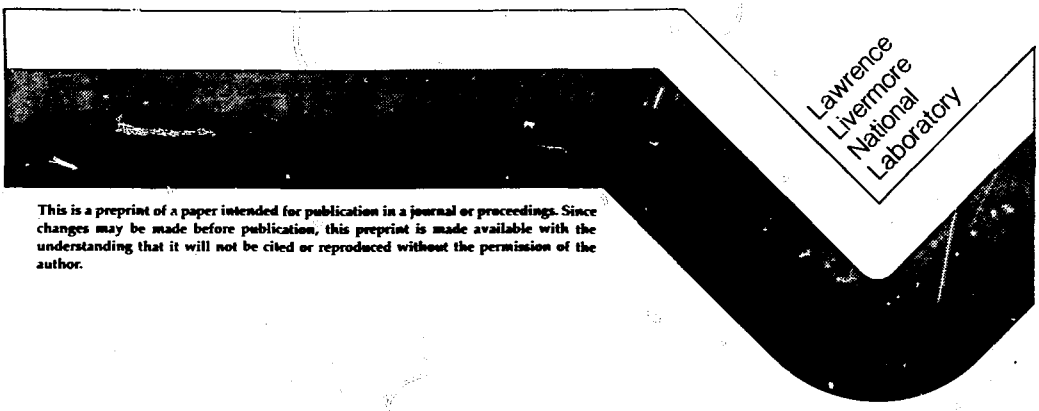
UCRL- 94163
PREPRINT

SAFEGUARDS RESOURCE MANAGEMENT

R. Scott Strait

This paper was prepared for submittal to
INMM 27th Annual Meeting
New Orleans, Louisiana
June 22-25, 1986



SAFEGUARDS RESOURCE MANAGEMENT

UCRL--94163

DE86 013223

R. Scott Strait

Lawrence Livermore National Laboratory*

Livermore, California

Abstract

Protecting nuclear materials is a challenging problem for facility managers. To counter the broad spectrum of potential threats, facility managers rely on diverse safeguards measures, including elements of physical protection, material control and accountability, and human reliability programs. Deciding how to upgrade safeguards systems involves difficult tradeoffs between increased protection and the costs and operational impact of protection measures. Effective allocation of safeguards and security resources requires a prioritization of system upgrades based on a relative measure of upgrade benefits to upgrade costs. Analytical tools are needed to help safeguards managers measure the relative benefits and costs and allocate their limited resources to achieve balanced, cost-effective protection against the full spectrum of threats. This paper presents a conceptual approach and quantitative model that have been developed by Lawrence Livermore National Laboratory to aid safeguards managers. The model is in the preliminary stages of implementation, and an effort is ongoing to make the approach and quantitative model available for general use. The model, which is designed to complement existing nuclear safeguards evaluation tools, incorporates a variety of factors and integrates information on the likelihood of potential threats, safeguards capabilities to defeat threats, and the relative consequences if safeguards fail. The model uses this information to provide an overall measure for comparing safeguards upgrade projects at a facility.

Introduction

Protecting nuclear materials is a complex problem for facility managers. Potential threats can vary from highly organized attacks by terrorist groups, which require skilled military-type responses, to small spontaneous acts of sabotage by disgruntled employees against which the normal physical safeguards may provide little protection. To counter the broad spectrum of potential threats, facility managers rely on diverse safeguards and security measures. However, deciding how to upgrade their safeguards systems involves difficult tradeoffs. Perhaps the greatest difficulty in determining the proper level of safeguards is deciding to what extent improved performance warrants increased cost. In general, the cost of operating facilities handling nuclear materials ultimately is borne by the public, which must be protected from excessive cost as well as excessive risk. Given the broad range of threats, the facility manager must decide where to commit additional resources, what improvements are justified, and how best to upgrade the system capabilities.

Effective allocation of safeguards resources requires a prioritization of system upgrades based on a relative site-wide measure of upgrade benefits to upgrade costs. The benefits of safeguards system upgrade generally can be expressed in terms of reduced risk. Risk reduction may result from increased deterrence, improved detection and prevention capabilities, or mitigation of potential consequences. Costs of the safeguards upgrades include both budgetary funds and operational impacts.

Due to the complicated nature of the problem and the large number of threats, targets, and safeguards alternatives, analytical tools are needed to help safeguards managers. These analytical tools should measure the relative benefits and costs of different safeguards upgrades on a site-wide basis and help to allocate the facility's limited resources to achieve balanced, cost-effective protection against the full spectrum of threats.

Lawrence Livermore National Laboratory (LLNL) has developed a conceptual approach and quantitative model to help safeguards managers evaluate alternative safeguards improvements at their facilities. This model incorporates all three aspects of the risk equation, consequences, capabilities, and likelihood, and takes a site-wide perspective. The model allows the safeguards manager to evaluate and compare the risk reduction of system upgrades that affect different adversar-

*Work performed under the auspices of the U.S. Department of Energy by the Lawrence Livermore National Laboratory under Contract W-7405-Eng-48.

ies (e.g., terrorists or insiders), threats (e.g., theft or sabotage) or targets (e.g., laboratory buildings or reactors). Only with a comprehensive model that incorporates all these different aspects of the problem can a manager develop unbiased and complete insight into the problem. A number of analytical tools have been developed to deal with limited aspects of the problem, but this is one of the few to take the important comprehensive approach. A facility manager can use the model results to (1) identify the threats to which the system is most vulnerable, (2) assess whether protection is adequate, and (3) determine which protection measures should be enhanced. The model is implemented on a portable computer, and it can be used "on the spot" to identify which expert judgments are most crucial.

This model is compatible with many existing, more detailed approaches to relevant subproblems. Inputs to the model can be derived from existing safeguards evaluation tools. Where such tools are not available, judgment may be elicited *from appropriate persons deemed authorities in* their particular subject areas. The opinions on each factor are then combined to produce quantitative statements of the overall risk at a given site.

The model is in a preliminary stage of implementation, and further development is underway. This paper presents the general approach used in the model and discusses its major components. The model, which is designed to complement existing nuclear safeguards evaluation tools, incorporates a variety of factors and integrates information on the likelihood of potential threats, safeguards capabilities to defeat threats, and the relative consequences if safeguards fail. The model uses this information to provide an overall measure for comparing the benefits of safeguards upgrades at a facility.

First the overall analytical approach is presented followed by sections on the identification of the adversaries, threats, and targets, the threat likelihood, the safeguards capabilities, and the consequences of failing to thwart a threat.

Analytical Approach

A good measure of the benefit of a safeguards upgrade is the reduction in risk achieved when the upgrade is implemented. Risk reduction should reflect all the effects of the improved safeguards, including increased deterrence, improved detection and prevention capabilities, and mitigation of potential consequences. In the model, we *compute the level of risk for a particular safe*guards system as the product of: the likelihood of a threat, the probability the threat will not be thwarted, and the expected consequences if the threat is not defeated.[1] These three elements of risk are discussed below.

The first step is to structure and scope the analysis. This is done by identifying and characterizing the potential adversaries (e.g., terrorists or insiders) threats (e.g., theft or sabotage), and targets (e.g., material or buildings). Regulatory guidance defines a minimum set of adversaries and threats of concern. However, the facility safeguards manager must also be sure to consider localized or site-specific adversaries. The threats considered need to be highly representative, and may include theft of different quantities and forms of nuclear materials and various types of sabotage. In addition to characterizing the types of potential adversaries and threats, the model user must define the targets against which these threats will be directed: e.g., particular buildings, operational systems, or blocks of buildings. The targets should be characterized so that all potential targets are included and so that the safeguards protecting an individual target are clearly defined.

Given the threats and targets identified in the first step, next the relative likelihood of each potential threat by adversary and target must be assessed. Because of the wide disparity in the probability of different adversaries, threats and targets, these relative likelihoods are key inputs to a site-wide evaluation of safeguards cost-effectiveness and the effectiveness of system upgrades. The assessment of the relative likelihoods for even a medium-sized facility can, unfortunately, be overwhelming due to the large number of combinations of adversaries, acts, and targets and the lack of information to support assessments. A significant contribution of this work is the development of a simple approach to this relative likelihood assessment based on a few reasonable assumptions about the adversaries' actions.

The next step is to determine the detection and prevention capabilities of the existing safeguards system. This is the ability of the system to thwart an adversary. This is measured by the probability of the adversary being defeated given that an attempt occurs. Most current safeguards evaluation methods address all or portions of this step. The use of these methods in the LLNL approach to safeguards resource management is discussed below.

The consequences of adversary success must then be determined. Consequences are measured on a relative scale, which combines such well-defined consequences as monetary costs with such elusive issues as national defense concerns. This measure of consequences can be assessed directly or by use of analytical models.

Each of the steps of the approach are discussed below. The final section discusses how increased deterrence, improved detection and prevention, and mitigation of consequences are reflected in this approach.

Identification of Adversaries, Threats and Targets

The evaluation approach begins with the identification of potential adversaries, threats and targets. This step establishes the scope of the analysis. The larger the number of adversaries, threats and targets identified, the more comprehensive the analysis. Naturally, the effort required to perform the analysis increases with the number of adversaries, targets and threats identified.

The U.S. Department of Energy provides guidance for its facilities on generic adversaries and threats that must be considered when designing safeguards and security systems.[2] In addition to

these generic threats, the facilities managers must consider site-specific threats that may present particular concerns to their facilities. At a minimum, the analyst would want to include terrorist groups, and a single insider or insiders in collusion. The list of adversaries may be expanded to include different sizes of terrorists groups, different numbers of insiders in collusion, or terrorist groups and insider(s) in collusion. Adversaries of concern may also include:

- Foreign governments
- Criminals
- Extremists
- Nuisances
- Different types of insiders.

For each adversary the threats posed by the adversary must be identified. These threats, which can be thought of as the adversary's goals, may include:

- Site penetration
- Radiological or industrial sabotage
- Theft or diversion of special nuclear material (SNM)
- Theft of a nuclear weapon or component
- Compromise of classified information
- Theft of government property.

Some of these threats may be defined more narrowly, such as by quantity of nuclear material. Some of these threats may not be applicable to all adversaries; for example, diversion of SNM (i.e., removing SNM from its authorized location, but not from the site) only makes sense for insiders.

For each adversary and threat, the analyst also needs to specify the potential targets within the facility. These targets may be defined by a physical location (e.g., reactor building) or by a specific system (e.g., computer network) and the adversary's goal. Targets with similar safeguards and similar consequences of successful threats may be grouped.

Assessment of Threat Likelihoods

The first element of the risk measure is the relative likelihood that a threat will occur. This is the element of the risk measure with the least available data to support the analysts' assessment. The likelihood needs to be assessed for each adversary, threat and target combination, which can be a very difficult task. The approach presented here divides this task into two parts, a simple basic assessment and a more difficult calculation based on assumptions about the adversaries' goals. By doing so, the threat likelihood assessment is made less burdensome and more defensible.

As mentioned above, the estimate of the likelihood of potential threats should be based on input data regarding potential adversaries, their primary motivations, their goals and resources, and the sites or targets they might choose. This step begins with the assessment of the relative likelihood of an attempt by each adversary (terrorists, demonstrators, insiders, etc.) by threat (theft of SNM, radiological sabotage, industrial sabotage, etc.) regardless of the specific target within the facility. The suggested approach is to first assess the relative overall likelihood of a threat for one adversary versus the others. Then, for each adversary, assess the relative likelihood of each type of threat by that adversary.

This assessment should, of course, involve experts in the field, e.g., representation from local and national law enforcement agencies, the national defense organizations, as well as experts on terrorism, foreign governments, criminal activity, and antinuclear extremist groups, human reliability programs and labor relations. In each case, questions should concern motivations, goals, resources, and targets of each of the adversaries.

The assessment should be structured and well documented. A suggested approach is to begin with the motivations of the potential adversaries. Motivations of adversaries may include:

- Monetary gain
- Embarrassment
- Extortion
- Stop nuclear power, weapons, or research
- Political power
- Nuclear weapon development
- Revenge
- Intelligence
- Pressure via blackmail.

Next, assess the relative likelihood that an adversary with a particular motivation would attempt to accomplish the objective at the facility. These assessments can then be combined in a computer model to produce a quantitative statement of the relative likelihood of each potential adversary. This approach is documented in Ref. 3. Based on the motivations of each adversary type, it is a somewhat smaller step to assess the relative likelihood of each threat by each adversary type.

The second part of this step of the threat likelihood assessment is to allocate the likelihood of an adversary attempt by threat to the different targets within a facility. For this allocation, a mathematical model is developed of the adversary's choice of targets. This model is based on a measure of the adversary's expected value for each target. This measure of value is probability that the adversary will succeed multiplied by the relative consequence of a successful act. This approach makes two basic assumptions:

1. The adversaries are rational and wish to maximize their expected gain.
2. The adversaries' gain from successful completion of their act is proportional to

the consequences of the threat to the facility.

The first of these assumptions, of course, may not be true and some adversaries may be irrational. However, if it is assumed that adversaries are indeed irrational then any method of predicting their targets becomes a meaningless exercise. The modification of this assumption to allow for risk-prone or risk-averse adversaries can be accomplished if it is deemed important. The second of these assumptions may not be valid, but it probably represents a good first approximation. Modification of the model to eliminate the need for this assumption is not particularly difficult (although the effort required for assessment of values may be quite large). Recall that the outputs of the model resulting from these assumptions are all relative and that their absolute values are not of concern and, in fact, may not be meaningful.

Using the adversary's expected value for each threat, the relative likelihood of each target is determined using a "logit" model. The logit model, commonly used in studies of consumer preference, bases the likelihood of adversary actions on the differences between relative expected values of alternative targets. Consequently, if two alternative targets have the same computed value, the adversary is assumed to be equally likely to choose either target. (For additional information on the logit model see "Conditional Logit Analysis of Qualitative Choice Behavior" by D. McFadden.)

The logit model calculates the probabilities of adversary target based on the following:

$$p(\text{adversary chooses target}) = \frac{\exp(a * \text{adversary's expected value of target})}{\sum_{\text{target}} \exp(a * \text{adversary's expected value of target})}$$

The constant "a" reflects the incomplete knowledge of the adversary's decision processes and imperfections in the model. The larger this constant, the more likely the adversary is to choose the target with the highest value, and the smaller this constant, the more uncertainty there is in the estimate of the adversaries' targets. One way to assess the value of the constant is by first assessing the difference in adversary expected values that results in two-to-one odds that the adversary will choose the target with the higher value. The value of the constant can then be determined by the following formula:

$$a = \frac{\ln(2)}{\text{difference in value}}$$

The threat likelihood assessment approach presented herein provides a logical framework for determining the relative likelihood of threats by adversary and target. As such, it provides needed input to a site-wide risk assessment. Since the model can be implemented on a portable computer, it can be used to assess expert judgments at various locations and to identify which expert judgments are most crucial in determining the relative likelihood of various threats.

Determination of Safeguards Capabilities

For each adversary, threat and target, the detection and prevention capabilities of the safeguards system need to be determined. These capabilities are measured as the probability that an adversary will be thwarted given the threat actually occurs. For example, the probability that a terrorist group will be stopped before it is able to leave the facility with any nuclear material, or the probability that an employee will be prevented from committing an act of sabotage. This determination may be made by direct assessment by knowledgeable individuals, e.g., the facility manager or security director, or by use of analytical model. One model that integrates the effects of safeguards for both the outsider and insider threats is the Method for Integrating SAVI and ET Results[5] (MI$ER) program recently developed for the DOE Office of Safeguards and Security by LLNL. This model employs the LLNL Safeguards Evaluation Method--Insider Threat[6] (ET) and the Sandia National Laboratories' Systematic Analysis of Vulnerability to Intrusion[7] (SAVI) for outsider evaluation. There are other analytical approaches, each with their own advantages and disadvantages.

Estimation of Consequences

The consequences of adversaries achieving their objectives must be estimated for each combination of adversary, threat and target identified in the first step of the analysis. The diverse nature of these threats requires that they be assessed using a relative subjective index. Consequences are very uncertain and can take many forms. For example, assessing the consequences of SNM theft range from damage to national security (e.g., use of the material in a nuclear weapon), to environmental and health effects (e.g., contamination of the atmosphere or drinking water). Of course, the consequences also include monetary costs such as the expense to recover or replace the material and the expenses of any clean-up.

Comparing diverse consequences of different threats requires difficult trade-offs. It is recommended that the analyst explicitly recognize these difficulties and incorporate uncertainties and trade-offs into a single relative index during the consequence assessment stage of the analysis. In assessing consequences, it is useful to construct probability (event) trees to represent the wide range of malevolent uses of nuclear materials, outcomes of sabotage, etc., and the injurious effects of those outcomes.

Since consequences are uncertain and their assessment is highly subjective, an analytic approach can improve the consistency of safeguards decisions by explicitly representing the uncertainties, value judgments, and trade-offs. Analysts can explore wide ranges of probability or value judgments on crucial variables. When coupled with other components of the resource allocation process, the consequence assessment stage can show how subjective judgments about consequences can affect specific safeguards decisions.

There are some analytic approaches to determine the expected consequences from low probability events. Applications of these techniques specifically to SNM are illustrated by Refs. 8-10.

Summary--Deterrence, Protection, and Mitigation

This model has been implemented on a microcomputer using a spreadsheet program. This implementation allows for the quick assessment of the importance of differences in judgments, the easy determination of the effectiveness of different safeguards upgrades, and the speedy reevaluation of upgrade priorities.

The model of adversary choice of targets is the key to reflecting the interrelated nature of deterrence, detection and prevention, and mitigation. Mitigation is the reduction in the loss should an adversary succeed in achieving their goals. This is particularly important for sabotage where, for example, duplicate capabilities may be constructed to mitigate industrial sabotage or redundant containment structures may be provided to mitigate radiological sabotage. The reduction in consequences is reflected in the overall relative risk, and the decrease in attractiveness of a target to the adversary is reflected via the calculation of target likelihood.

Improvements in the detection and prevention capabilities of the facility are directly incorporated in the overall measure of facility relative risk, but improvements in the safeguards of specific high risk facilities will result in lower likelihoods of adversaries targeting those facilities. This secondary result occurs because of the method of calculating the threat likelihoods by target.

Overall deterrence of adversaries can be reflected in the preliminary assessments of adversary and threat likelihoods. This may require additional judgments from experts. However, due to the two-step procedure for computing threat likelihood, the model can recompute target likelihood given the change in judgment regarding adversary and threat likelihood.

The model presented above is well-suited for allocating safeguards resources on a facility-wide basis by providing a single relative measure of the risk to the facility under alternative safeguards systems. The design of the model is particularly advantageous for repetitive evaluations of different safeguards upgrades, because it automatically incorporates the interrelated benefits of deterrence, prevention, and mitigation of different safeguards. By comparing the relative improvements in the index of facility risk for different systems with the cost of those systems, the most cost-effective system can be determined and the facility safeguards resources allocated accordingly.

References

1. Bennett, C. A., W. M. Murphy, and T. S. Sherr, "Societal Risk Approach to Safeguards Design and Evaluation," U.S. Energy Research and Development Administration, ERDA-7, June 1975.

2. Rose, R., "Generic Threats for DOE Nuclear Programs and Facilities," U.S. Department of Energy, Washington, DC, Memorandum, January 31, 1983.

3. West, D. J., R. A. Al-Ayat, and B. R. Judd, "DOE Site-Specific Threat Assessment," Lawrence Livermore National Laboratory, UCRL-53668, July 1985.

4. McFadden, D., "Conditional Logit Analysis of Qualitative Choice Behavior," in Frontiers in Economics, edited by P. Zarembka, Academic Press, New York, 1973.

5. "Method for Integrating SAVI and EI Results, the MI$ER Computer Program Users Manual," Lawrence Livermore National Laboratory, June 1986.

6. Safeguards Evaluation Method--Insider Threat, Workbook, Lawrence Livermore National Laboratory, 1984.

7. "SAVI--Systematic Analysis of Vulnerability to Intrusion, Users Guide," Sandia National Laboratories, 1986.

8. Hill, G. B., "Societal Consequences of Malevolent Situations: Implications for Safeguards Policy," Decision Science Consortium, Inc., Technical Report 81-8, May 1982.

9. Crane, F. L., B. W. Welles, and J. E. Owens, "Consequence Risk Analysis of Sabotage at DOE Facilities, Phase I," International Energy Associates Limited, IEAL-181, December 1980.

10. Indusi, J. P. and A. Fainberg, "A Summary of the Estimated Consequences of Malevolent Acts Involving Nuclear Material and DOE Facilities," Brookhaven National Laboratory, 1981.