MASTER

SAND79-0059C
NUREG/CR-0788

Safeguards Methodology Development History*

L. D. Chapman, H. A. Bennett, D. Engi,
L. M. Grady, B. L. Hulme and D. W. Sasser
Sandia Laboratories, Albuquerque, New Mexico, USA

ABSTRACT

The development of models for the evaluation and design of fixed-site nuclear facility, physical protection systems was under way in 1974 at Sandia Laboratories and has continued to the present. A history of the evolution of these models and the model descriptions are presented. Several models have been and are continuing to be applied to evaluate and design facility protection systems.

SUMMARY

The development of models to aid in the evaluation of physical protection systems of nuclear facilities began at Sandia Laboratories as early as 1974. This work has been sponsored principally by the United States Nuclear Regulatory Commission. The purpose in developing these models is to fulfill the need for:

1. A consistent approach to the evaluation of the effectiveness of physical protection systems in defending against a hypothesized adversary threat, and
2. A quantitative technique for determining upgrades to existent facilities and for designing new facilities.

Two of the first physical protection models developed were the Forcibile Entry Safeguards Effectiveness Model (FESEM) and the Insider Safeguards Effectiveness Model (ISEM). These are Monte-Carlo simulation

*This report was prepared for the U.S. Nuclear Regulatory Commission, Office of Nuclear Regulatory Research under research order No. 60-78-1090 and under the auspices of the U.S. DOE, contract No. 40-550-75.

programs which consider a single target and a single adversary path and include simple engagement models. Through experience gained in applying these techniques, needs for other models were identified.

The simplistic representation of adversary and guard tactics incorporated into FESEM and ISEM lacked the ability to define strategies and to incorporate other pertinent details. The Fixed-Site Neutralization Model was developed to provide detailed interactions of individuals, consideration of individuals' abilities and skills, and leader planning capabilities which represent the neutralization process. The Safeguards Network Analysis Procedure was developed to provide a capability to represent user-specified interactions of individuals.

The large run times and difficulty in performing sensitivity analyses with the early simulation codes made the evaluation of all adversary scenarios infeasible. This spawned the development of analytic evaluation techniques which could facilitate total system evaluations. Target identification, critical path analysis and path evaluation were combined into an overall model called the Safeguards Automated Facility Evaluation. This model begins with a facility layout on which significant barriers and targets have been identified. The analyst digitizes the facility blueprints into the form of a network of nodes and arcs. Pathfinding algorithms are utilized to find critical adversary paths and analyses of these paths are then conducted. This model is user-oriented and makes use of interactive computer timesyaring and computer graphics.

EARLY SCENARIO BASED MODELS

Two of the first safeguards evaluation models which were developed were the Forcible Entry Safeguards Effectiveness Model (FESEM)[1] and the Insider Safeguards Effectiveness Model (ISEM)[2]. FESEM and ISEM employ Monte-Carlo techniques to simulate a group of adversaries attacking a nuclear facility. The principle difference between these two models lies in the hypothesized threat that they are structured to address. FESEM was structured to consider primarily adversaries who do not have authorized access to the facility (outsiders) while ISEM focuses on adversaries who do have authorized access (insiders).

The focus of each of the models is defined in terms of the hypothesized threat (outsiders vs. insiders), and their internal structures reflect that difference in emphasis. For example, the neutralization (battle) submodel within FESEM can accommodate any number of adversaries. In contrast, although ISEM can consider any number of adversaries who might covertly tamper with the detection system, its neutralization submodel can accommodate only one adversary who can become engaged in combat with the security force.

Experience gained through the application of FESEM and ISEM provided the impetus for further safeguards methodology development. There were essentially two schools of thought regarding the most fruitful direction for further developmental work. On the one hand it was clear that the single-scenario orientation of FESEM and ISEM was not amenable to an evaluation of safeguards systems considered in their entirety. That is,

an evaluation of the effectiveness of a safeguards system in countering individual adversary scenarios merely reflects the system's ability (or inability) to deal with those scenarios - it is likely to imply little of the safeguards system as a whole. Consequently, a need for a global approach to the problem of evaluating safeguards system effectiveness was identified.

At the other extreme, both FESEM and ISEM were criticized for not including a sufficient amount of detail in individual scenarios. This criticism was directed primarily at their inability to represent complex tactics that might be used by the adversaries as well as the security force.

To satisfy both of these concerns, developmental activities proceeded along two lines. One area of work centered on the development of detailed scenario models. This work resulted in a second generation of scenario models that can explicitly represent quite complex tactics. The other area of work focused on developing a global approach to safeguards effectiveness evaluation. The result of the global effort is an interlinked collection of analytical techniques which can be used to evaluate the effectiveness of the entire safeguards system in terms of a piece-wise evaluation of the disaggregated system. This disaggregation allows for a significant simplification which facilitates a global treatment of the problem. The next two sections describe in more detail the products of these two developmental activities.

SECOND GENERATION SCENARIO MODELS

The primary thrust in the development of the second generation scenario models was in the direction of enhancing the capability to represent complex tactics. This enhanced capability was pursued through the development of two separate scenario models. One of these models, the Fixed-Site Neutralization Model (FSNM)[3], was developed with the intent of representing tactics internally in the model's logic with a minimal amount of user input of a tactical nature. The other scenario based model, the Safeguards Network Analysis Procedure (SNAP)[4], is the antithesis of FSNM with respect to the representation of tactics. SNAP requires explicit user input to represent the tactical process. Both models employ Monte-Carlo techniques to simulate randomness in the scenario. Outputs from the models include estimates for a variety of system performance measures.

Fixed-Site Neutralization Model (FSNM)

FSNM consists of a representation of the facility and personnel along with their activities and decision processes. The facility is represented in the model as a rectangular area which may, and probably should, extend beyond the boundaries of the actual facility.

Architectural features of the facility, such as buildings, fences, walls, and outside areas (yards) are represented, together with interior features such as roofs, floors, stairs, doors, rooms, and halls. Such details as the visibility through a barrier, the difficulty of penetrating the barrier, and whether a door is closed or open, locked or unlocked, are explicitly modeled. The locations of sensors and their types, coverages, and operational states also appear in the facility description. The goals of the adversary are represented by specifying locations in the facility which are required to be occupied by some number of adversaries, possessing certain equipment, for a certain length of time.

Individual persons in the model, called "players", are represented in considerable detail. The representation has three aspects: physical, potential, and psychological. Adjustment of any or all of these aspects permits the simulation of differences between individuals or forces due to training, ability, or equipment. The physical aspect of a player's representation includes his location, posture, weapons and equipment, and physical status. The weapons and equipment a player carries may include pistols, rifles, grenades, light antitank weapons, ladders, keys, and other equipment. The characteristics of each type of weapon, including range, ammunition supply, and effectiveness against various targets, are represented.

Players in the model have three main activities in which they may decide to engage during a simulation time period. These activities are to move, fire, or observe. Other activities may also occur, including surrendering to or capturing an opponent. Every player has an associated collection of perceptions about observable entities (people, vehicles, and sensors) at the facility. These perceptions form, in effect, the "memory" of a player and may change as the result of direct observation by the player or by his reception of information from other players over communications systems.

Safeguards Network Analysis Procedure (SNAP)

SNAP is a simulation language developed specifically for modeling safeguards systems. With the SNAP approach, the analyst constructs a model of the safeguards system by interconnecting a set of SNAP symbols to represent the system elements and their interactions. The resulting SNAP networks are then transferred to a computer compatible form by data cards representing the symbols and their interconnections.

Using the SNAP procedure for safeguards modeling, one combines knowledge of the system, scenarios, modeling objectives, and the SNAP symbology to develop a network model of the system under consideration.

This network model is a graphic representation of the nuclear facility, guard operating policies and adversary attack scenario. Typically, the elements of this network model will form a one-to-one correspondence with the components of the actual physical system and scenario being studied. Due to this relationship, a SNAP network provides an excellent communications vehicle. SNAP symbols have been designed to represent the individual elements of a nuclear safeguards system, thus the translation from a system element to the SNAP symbol should be direct.

A SNAP network model is composed of the facility subnetwork, the guard subnetwork, and the adversary subnetwork which interact to produce the overall behavior of the safeguards system. Items which flow through network models are referred to as transactions. The transactions which flow through a SNAP network are guard forces and adversary forces. The force is the most fundamental level of detail in SNAP and represents one or more individuals acting as a single unit.

The facility subnetwork is the most fundamental. It is a static network in the sense that transactions do not flow through it during the simulation. Its purpose is to define the various elements of the facility and their relationships. These elements may include fences, yards, nuclear material, storage vaults, doorways, rooms, sensors, etc. The guard subnetwork defines guard operating policies and includes a representation of the guard decision logic as well as their physical movement through the facility. Guard forces are the transactions which flow through the guard subnetwork. The adversary subnetwork is treated in a similar manner.

The flexibility afforded by SNAP makes it the preferred approach to modeling scenarios. In effect, all of the modeling capabilities of FESEM and ISEM are included in SNAP. Moreover, if a sufficient amount of detail is incorporated into the facility, adversary, and guard submodels, the level of resolution can be equal to that of FSNM. It is worth noting that the inherent flexibility of SNAP is a result of the modeling philosophy used in its development. That is, the SNAP analysis program can be viewed as a simulation "language" specially tailored to model safeguards scenarios.

With the advent of SNAP, the majority of the criticism directed at the limitations of the early scenario models (FESEM and ISEM) were answered. SNAP can be used to represent quite complex

tactical situations and, as a consequence, lends credibility to the evaluation of individual scenarios. In the context of "vulnerability analyses", SNAP is a valuable tool in that it can provide insights into the strengths (or weaknesses) of the safeguards system in defending against a predefined adversary scenario. However, as previously observed, the analysis of a single scenario is likely to offer little in the way of global insights with respect to the safeguards system. Moreover, even without considering analyst time, a detailed analysis of a sufficient number of scenarios to gain these global insights is unlikely to be computationally tractable. In addition, it is not obvious just what is implied by "a sufficient number of scenarios". To address these inherent limitations which are inexorably linked to any scenario based technique, a global approach to the evaluation of safeguards effectiveness was developed.

A GLOBAL EVALUATION MODEL

The principle limitations of the scenario based models with respect to their applicability to a global safeguards effectiveness evaluation were observed to be of a philosophical as well as a technical nature. First, on the technical front, the scenario based models involve relatively complex Monte-Carlo simulation techniques. In addition to the significant amount of computer time necessary to replicate a sufficient number of times to obtain statistical stability, the time required of the safeguards analyst in preparation of the input for a single scenario can be excessive. Perhaps more importantly, the modeling philosophy of the scenario based models does not include the "generation" of adversary scenarios.

The Safeguards Automated Facility Evaluation (SAFE)[5] methodology evolved as a result of efforts to overcome the limitations described above. The technical limitations were addressed by developing a set of analytical techniques which are computer-time efficient and by structuring a highly user-oriented approach that is analyst-time efficient. On the philosophical level, techniques for generating "optimal" adversary scenarios were developed.

SAFE consists of a collection of functional modules for facility representation, component selection, adversary path analysis, and effectiveness evaluation. SAFE combines these modules into a con-

tinuous stream of operations. The technique has been implemented on an interactive computer time sharing system and makes use of computer graphics for the processing and presentation of information. Using this technique, a global evaluation of a safeguards system can be provided by systematically varying the parameters that characterize the physical protection components of a facility to reflect the perceived adversary attributes and strategy, environmental conditions, and site operational conditions.

The SAFE procedure requires as input, a blueprint of the facility, showing the facility layout characteristics, the targets, and vital areas. To obtain this input, the analyst must perform a facility characterization activity[6]. Relevant sources of information for this activity include the security plans, facility drawings, safety analysis reports, environmental reports, and site visits. Based upon this information, the analyst must synthesize the necessary facility layout characteristics, targets and vital areas, operational conditions, site-relevant environmental conditions, physical protection system and guard characteristics for which analyses are to be performed.

The first step in the application of SAFE is to construct a computer representation of the facility. This representation provides an explicit record of the analysts assumptions concerning the facility. For example, the analyst would indicate all principle barriers and obstacles to adversary movement, all points of potential ingress and egress, floor levels and interconnections, and targets and vital areas for specific operational conditions. This information is used to organize and digitize the pertinent facility information into a computer usable form. The final output of the facility representation is a graph in which nodes represent potential access points or targets, and arcs represent possible movement between nodes.

The next phase in the SAFE analysis requires the analyst to set component performance for individual safeguards elements. The specific performance for both hardware and personnel "components" should be based upon relevant sets of environmental and adversary conditions. The analyst uses the component performance to determine weights for all nodes and arcs in terms of detection probabilities and time delays for adversary penetrations. Appropriate selection of these weights provides bounds for a range of adversary attributes. The resultant graph-theoretic representation serves as input to the adversary path analysis module within SAFE.

The generation of adversary scenarios is achieved by selecting optimal paths through the facility for the adversary. Both

theft and sabotage path selection were previously accomplished by several alternative techniques[7,8,9,10]. Currently, SAFE uses one of three measures for adversary pathfinding: 1) minimum adversary task time, 2) minimum adversary detection probability, and 3) minimum timely-detection of the adversary. Within SAFE, these measures can be either deterministic[11] or stochastic[12]. In effect, the timely-detection method generates paths which minimize the probability that the security force can confront (or interrupt) the adversary. The output of the adversary path analysis is a collection of ordered sets of node identifiers that represent physical paths in the facility which are the most "critical" in terms of the measure being used. This information is a portion of the input to the effectiveness evaluation module in SAFE.

Effectiveness evaluation can be decomposed into two major parts: interruption and neutralization for a given path. The path is "evaluated" by first determining the probability that the adversary will be interrupted and then determining the probability that the adversary will be neutralized or defeated by the security force. These two probabilities can be multiplied together to yield the total probability that the physical protection system will be successful in defending against the adversary along the path under consideration.

The Estimate of Adversary Sequence Interruption (EASI)[13] model is an analytical technique which is used in the effectiveness evaluation module to compute the probability that the adversary will be interrupted. EASI focuses on the adversary path and requires information related to the probability of detecting the adversary, the time required for determining the proper response, the probability of communication with the security forces, the delay along the adversary path and the response time of the security force. The output of EASI is an estimate of the probability of adversary interruption along the specified path, i.e., the probability that the security force arrives at a point along the adversary's path prior to the time that the adversary passes through that point.

The Brief Adversary Threat Loss Estimator (BATLE)[14] model is an analytical technique that is used to estimate the probability that the adversary is neutralized by the security force. In addition to the distance between combatants, the information required by BATLE is the type of weapons, the recency of training, the amount of cover, and the number and timing of arrivals of reinforcements for the adversary as well as the security officers. The output of BATLE is the probability that the adversary is neutralized by the security force. This "neutralization probability" is then multiplied by the "interruption probability" to yield the total probability of success of the physical protection system for the path in question.

Capabilities for effectiveness evaluation can be utilized in either a single or multipath mode. During a single path evaluation using EASI, the probability of interruption is calculated and the user may request two- or three-dimensional plots which show the probability of the adversary interruption as a function of one or two of the other input variables[15]. Based on the probability of interruption, these graphs illustrate sensitivities related to upgrading the facility. The multipath option displays in tabular form the probability of interruption, the traversal time of each path, and the frequency at which nodes appear in the set of critical paths. The multipath evaluation identifies paths that are particularly vulnerable and thus are candidates for study by more elaborate evaluation modules such as the scenario based models previously described.

COMMENTARY

Generally, the scope of a safeguards evaluation model can effectively address one of two issues:

1) global safeguards effectiveness, or
2) vulnerability analysis for individual scenarios

SAFE addresses 1) in that it considers the entire facility; i.e, the composite system of hardware and human components, in one "global" analysis. SNAP addresses 2) by providing a safeguards modeling symbology sufficiently flexible to represent quite complex scenarios from the standpoint of hardware interfaces with other elements of the physical protection system while also accounting for a rich variety of human decision making.

A combined SAFE/SNAP approach to the problem of safeguards evaluation logically proceeds along the following lines:

1) Initially, apply SAFE to identify global safeguards vulnerabilities,
2) Represent these vulnerabilities in scenarios that can be analyzed using SNAP.

Conceivably, insights of a global nature (especially as they relate to guard tactics and deployment strategies) could be gained from the SNAP vulnerability analysis. These insights might be formally "fedback" to SAFE, thus closing the global/scenario evaluation loop.

It should be emphasized that the safeguards analyst should remain intimately involved with the evaluation at every stage. Due to the complexity of safeguards problems, information gained by exercising the evaluation models described herein is intended to be of a supplementary nature only. That is, the analyst should consider the output of the models as inputs to the holistic evaluative process.

References

1. Chapman, L. D., et al., "Users Guide for Evaluating Alternative Fixed-Site Physical Protection Systems Using 'FESEM'" SAND77-1367, Sandia Laboratories, Albuquerque, NM, November 1977.

2. Boozer, D. D., Engi, D., "Insider Safeguards Effectiveness Model (ISEM) Users Guide," SAND77-0043, Sandia Laboratories, Albuquerque, NM, November 1977.

3. Engi, D., et al., "Fixed-Site Neutralization Model, Volume I, Executive Summary," SAND79-0063, Sandia Laboratories, Albuquerque, NM, January 1979.

4. Grant, F. H., Miner, R. J., Engi, D., "A Network Modeling and Analysis Technique for the Evaluation of Nuclear Safeguards Effectiveness," NUREG/CR-0616, Sandia Laboratories, NM, December 1978.

5. Chapman, L. D., et al., "Safeguards Automated Facility Evaluation (SAFE) Methodology,", SAND78-0378, Sandia Laboratories, Albuquerque, NM, August 1978.

6. Varnado, G. B., et al., "Reactor Safeguards System Assessment and Design, Volume I," SAND77-0644, Sandia Laboratories, Albuquerque, NM, June 1978.

7. Hulme, B. L., "Graph Theoretic Models of Theft Problems. I. The Basic Theft Model," SAND75-0595, Sandia Laboratories, Albuquerque, NM, November 1975.

8. Hulme, B. L., "Pathfinding in Graph-Theoretic Sabotage Models. I. Simultaneous Attack by Several Teams," SAND76-0314, Sandia Laboratories, Albuquerque, NM, July 1976.

9. Hulme, B. L., Holdridge, D. B., "SPTH3: A Subroutine for Finding Shortest Sabotage Paths," SAND77-1060, Sandia Laboratories, Albuquerque, NM, July 1977.

10. Hulme, B. L., Holdridge, D. B., "KSPTH: A Subroutine for the K Shortest Paths in a Sabotage Graph," SAND77-1165, Sandia Laboratories, Albuquerque, NM, August 1977.

11. Hulme, B. L., "MINDPT: A Code for Minimizing Detection Probability Up To a Given Time Away From a Sabotage Target," SAND77-2039, Sandia Laboratories, Albuquerque, NM, December 1977.

12. Engi, D., Shanken, J. S., "PATHfinding Simulation (PATHS) User's Guide," SAND78-2177, Sandia Laboratories, Albuquerque, NM, to be published.

References (Continued)

13. Bennett, H. A., "User's Guide for Evaluating Physical Security Capabilities of Nuclear Facilities by the EASI Method," SAND77-0082, Sandia Laboratories, Albuquerque, NM, June 1977.

14. Engi, D., Shanken, J. S., "Brief Adversary Threat Loss Estimator (BATLE) User's Guide," SAND78-1136, Sandia Laboratories, Albuquerque, NM, to be published.

15. Sasser, D. W., "User's Guide for EASI GRAPHICS," SAND78-0112, Sandia Laboratories, Albuquerque, NM, March 1978.