NUREG/CR-1140 Vol. II
UCRL-52712 Vol. II

AGGREGATED SYSTEMS MODEL OF NUCLEAR SAFEGUARDS

Prepared for
U. S. Nuclear Regulatory Commission
by
Lawrence Livermore Laboratory

Available from U. S. Nuclear Regulatory Commission Washington, D. C. 20555

Available from National Technical Information Service Springfield, Virginia 22161

NUREG/CR-1140 Vol. II
UCRL-52712 Vol. II

AGGREGATED SYSTEMS MODEL OF NUCLEAR SAFEGUARDS

Manuscript Submitted: June 28, 1979
Date Published: February 1980

Prepared for
Research Division
U.S. Nuclear Regulatory Commission
Washington, D. C. 20555
Under Interagency Agreement DOE 40-550-75
NRC FIN No. A-0115
by
Lawrence Livermore Laboratory
Livermore, CA 94550

operated by University of California

for the U.S. Department of Energy

FOREWORD

This two-volume report describes the Aggregated Systems Model (ASM), a formal aid for nuclear safeguards decision making. This tool permits decision makers to integrate various forms of safeguards information (adversary characteristics, safeguard system effectiveness, costs of safeguarding, consequences of diverted special nuclear material) to provide an evaluation and ranking of complex safeguards alternatives.

The work reported here had its origins in studies for the Nuclear Regulatory Commission begun in 1977 at LLL by John Lathrop, Stein Weissenberger, and Ivan Sacks, with subcontractors Bruce Judd of Applied Decision Analysis, Inc. (ADA) and Rex Brown of Decisions and Designs Inc. During this period, the basic ideas and structure were worked out for a highly aggregated model of safeguards decision making, to provide a tool for organizing the analysis of this very complex problem.

In 1978, the concepts and models were further developed by Weissenberger at LLL and Bruce Judd and Jean Huntsman of ADA; the major thrust of this period was to refine the model and examine value-impact tradeoffs in evaluating and ranking decision alternatives. The most recent effort, which has added significant features and which forms the immediate substance of the work reported here, has been carried out by Rokaya Al-Ayat of LLL and Judd and Huntsmann of ADA.

This report consists of two volumes. The first volume--Executive Summary--summarizes the methodology and introduces some of the results that have been achieved. The second volume describes in detail the Aggregated Systems Model.

Stein Weissenberger

CONTENTS

FIGURES

TABLES

ABSTRACT

When setting the performance criteria for systems that safeguard special nuclear material (SNM), decision makers must consider characteristics of the adversaries who attempt to divert SNM, safeguards responses to these attempts, costs of safeguards systems, and the consequences of diverted SNM.

This report describes an Aggregated Systems Model that is designed to assist decision makers in integrating and evaluating these diverse factors consistently. The report summarizes the results obtained from applying the model to safeguards decision making in areas such as the hardware and procedures required, substitution of electronics for human safeguards, and overall performance criteria for safeguards systems. New performance criteria designed to measure how safeguards systems deter adversary attempts are also described.

I. AGGREGATED SYSTEMS MODEL DESCRIPTION

INTRODUCTION

Background

The Nuclear Regulatory Commission (NRC) is responsible for protecting the public against malevolent uses of special nuclear material (SNM).[1] The Lawrence Livermore Laboratory (LLL) is developing analytical procedures to assist the NRC in meeting this objective. These procedures will assist in two areas:

- Setting safeguards criteria for facilities handling SNM
- Assessing individual plants to determine whether these criteria are satisfied.

This report describes the Aggregated Systems Model (ASM) developed to aid the NRC in both areas of responsibility. Under the general title of setting safeguards criteria, we consider the following kinds of NRC decisions (arranged in order from limited to global scope):

- Component or procedure requirements: Specifying the hardware and/or procedures to be followed in a facility
- Integrated safeguards evaluation: making tradeoffs among safeguards components with different functions (e.g., material control versus physical security) to achieve the most effective safeguards system, possibly within cost constraints
- Performance-based regulations: choosing performance criteria for individual components or for overall safeguards systems
- Value-impact analysis: finding the socially optimal safeguards system with the best balance between the social need for safeguards and the social cost of the safeguards system.

The analytical models described here help assess individual plants in two ways:

- Preliminary assessments: making a first-cut assessment, using subjective data and the ASM, which will highlight crucial areas and hence guide resource allocations in detailed analyses

- Summarizing detailed assessments: aggregating the results from a facility's detailed assessment to test sensitivity to changing parameters and to facilitate communication among regulators, operators, and designers.

Definition

The Aggregated Systems Model is a comprehensive quantitative tool for analyzing the types of safeguards decisions described in the opening paragraphs. It covers many diverse factors in these decisions, including adversary threats, safeguards responses, and the consequences of successful diversions. It represents, at an aggregate level, information that is usually developed by separate detailed analyses of these diverse factors. However, it contains only a small fraction of the data used by the detailed analyses. For a summary of other definitions used here, see the Glossary at the end of the report.

Report Organization

The remainder of Section I describes the analytical forms of the Aggregated Systems Model, including detailed descriptions of the adversary-facility interaction during an attempt. Section II describes data inputs for the illustrative analysis contained in Section III. Section IV gives conclusions and directions for further study.

AGGREGATED SYSTEMS MODEL (ASM) OVERVIEW

Overview Figure

The major elements of the Aggregated Systems Model are shown in Fig. 1. The Diversion Model contains data that characterize adversaries, their attempts, and the facility's safeguards system's response to those attempts. We frequently use the name "Diversion Model" to mean the combination of the Adversary and Facility models. The Adversary Model includes adversary characteristics and the adversary's choice of strategy and tactics. The Facility Model represents the safeguards system's ability to detect and interrupt the adversary's attempt.

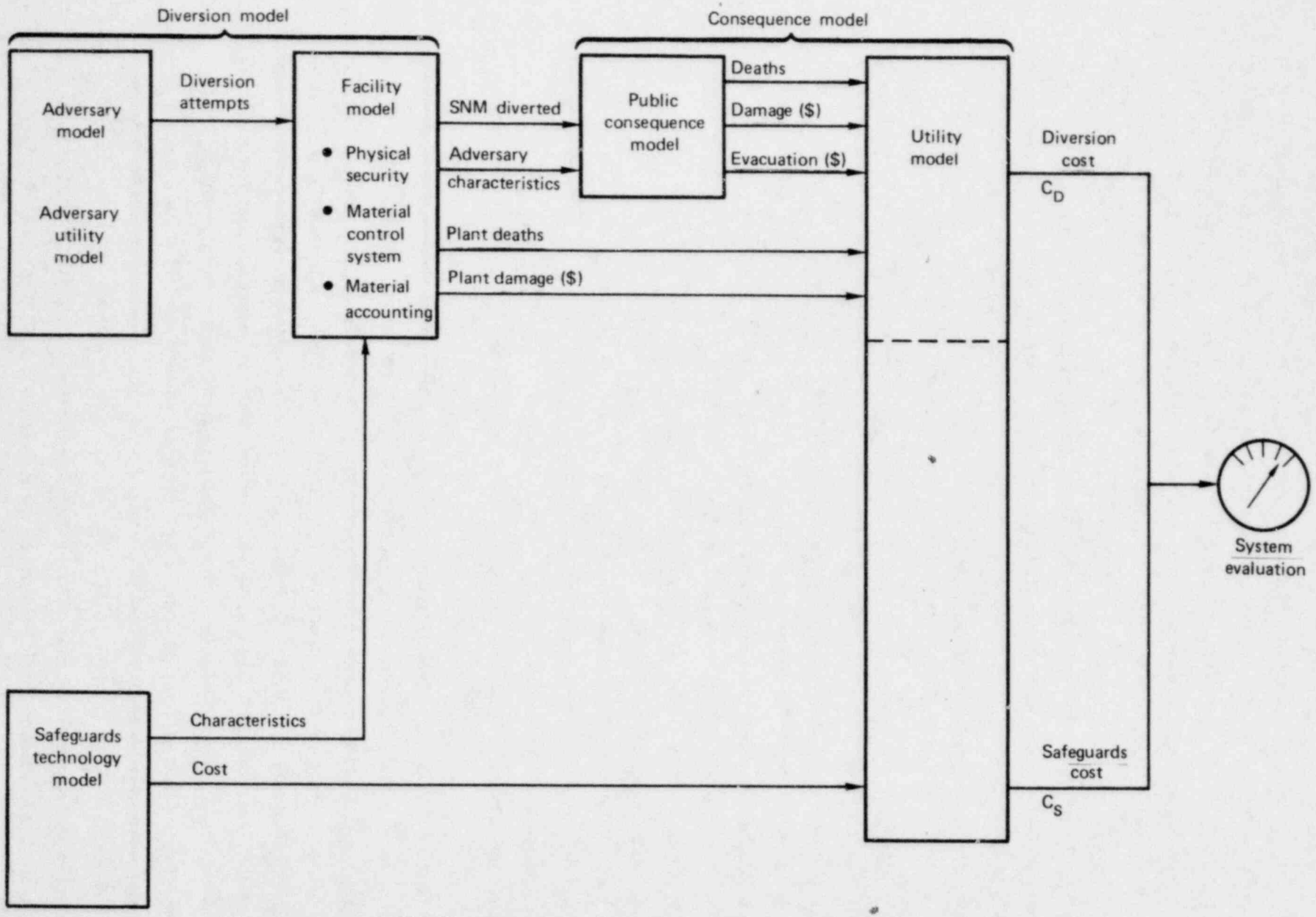


FIG. 1. Aggregated systems model overview.

The Consequence Model includes Public Consequence and Utility submodels. The Public Consequence Model describes possible malevolent uses of stolen SNM and the consequences. The Utility Model represents the decision-makers' assignment of societal values to all possible outcomes and the explicit tradeoff between safeguards costs and the risks of diverted SNM.

While the Diversion and Consequence Models quantify diversion risk, the Safeguards Technology Model quantifies economic costs of safeguards. The meter at the right-hand side of Fig. 1 shows the combined evaluation of safeguards costs and diversion risk.

Using the ASM

Recall the ASM uses listed at the beginning of this chapter:

- Setting safeguards criteria
  - Component or procedure requirements
  - Integrated safeguards evaluation
  - Performance-based regulation
  - Value-impact analysis
- Aiding detailed assessments
  - Preliminary assessment
  - Summarizing results.

We can conceptualize how the ASM will be used for each of these functions.

Components or procedure requirements and performance-based regulation of safeguards systems require only a portion of the ASM: the Adversary, Facility, and Safeguards Technology models. The regulator can change the list of required components or procedures, and then the ASM will show how overall system performance changes. Or the regulator can change a performance-based regulation, and the ASM will show whether or not safeguards components in the system can meet the regulation. Examples of performance requirements are: probability of detection, probability of interruption, limitations on the expected amount of SNM diverted each year, etc.

Integrated safeguards decisions use the same three models. In this case, the ASM is used to compare and to choose combinations of different safeguards (such

as physical security forces, material control systems, or accounting systems) that achieve a given level of performance. The model can show different ways to meet the performance requirement, some of which may be more advantageous to the operator than others. For instance, one could trade off improved real-time accounting equipment with frequent plant inventory shutdowns. Alternatively, one could observe the change in plant performance given new rules for upgrading safeguards.

Value-impact studies are designed to help decision-makers evaluate the social benefits and costs of changing safeguards requirements or system designs. The evaluation has three steps. The first step is analyzing how the designs or requirements affect the facility's performance against adversaries. In Fig. 1, this analysis is accomplished in the box labeled Facility Model. Performance can be measured by the arrow labeled "SNM diverted," which emanates from the Facility Model. The second step is to express the public consequences of the given performance level. For example, the ASM can describe how a tightening of security reduces the amount of SNM diverted and hence reduces the potential number of deaths or damage due to malevolent use of SNM. This translation occurs in the Consequence Model box shown in Fig. 1. The third step involves balancing changes in public consequences (values) with the social costs (impacts) of the safeguards to achieve the changed consequences. The Utility Model can help in this tradeoff process. However, if the decision-maker feels comfortable with implicit (rather than explicit) tradeoffs, the Utility Model need not be used for value-impact decisions. Notice that the value-impact analyses with the ASM use more of the models in Fig. 1 than do other uses.

As an aid to detailed assessments, the ASM can use some or all of the models in Fig. 1. In preliminary assessments, the model needs at least the Safeguards Technology and Facility Models, especially when evaluating performance against a specified adversary threat. The Adversary Model is used to consider the entire range of threats, rather than only one. The output of a preliminary assessment is a statement of the given facility's performance, measured as probabilities of detection, interruption, or expected SNM diverted.

When the ASM is used to summarize detailed assessments, more of the models may be used. If the goal is to show which facility components are the weakest

links in safeguards, only the Facility and Adversary Models are needed. At the other extreme, a statement of the total social risk posed by the plant may require all of the models in Fig. 1.

Assumptions in the Analysis

We conclude this overview with several assumptions that should be stated before discussing the model details.

- The data are purely illustrative.
- Only diversion attempts are modeled; sabotage is not.
- The hypothetical facility is the test bed design.[2]
- Consequences are limited to terrorist acts, and do not include international nuclear proliferation.
- The societal utility of lives, damage, and dollars is traded off in a linear fashion; the overall utility function for society is nonlinear; it is a step function dependent on amounts of SNM diverted.

Another important assumption is that the adversary choice is based on his or her expected utility. An alternative approach is to assume adversary choices based on other criteria, such as minimizing the probability of detection, maximizing the probability of success, etc. Sensitivity of the results to some of these alternative approaches is discussed in Section 3.

An additional major assumption is the probabilistic nature of the repeated events model. We assume that the probabilities of acquiring SNM, and of being detected, identified, or interrupted, do not change from one attempt to the next, unless the adversary is identified. This allows us to construct a tractable Markov model to represent the dynamic and probabilistic nature of repeated attempts.

Finally, we assume that the adversary preselects the diversion strategy and tactics before the first attempt, and this decision is not changed during the attempts. This understates the diversion risk somewhat, but we do not feel that the understatement is significant. This "preselection" assumption greatly simplifies the dynamic programming problem posed by the repeated attempts.

FACILITY MODEL

The Facility Model represents the system designed to safeguard SNM. There are nine basic building blocks, called "components," in the current version of the ASM:

- Quantity estimators
- Process state monitors
- Personnel monitors
- Procedure monitors
- Stationary guards
- Roving guards
- Two-person rule
- Nominal accounting system
- Frequent physical inventories

(Note that each of these building blocks is actually an aggregated system, composed of many hardware or software items.)

When referring to the safeguards system "design," we mean the list of components installed in the facility. Changing the design means adding, deleting, or substituting components.

The performance of the safeguards system is determined by the performance of individual components. Calculating system performance requires input data on component performance, and an algorithm for aggregating component performance to determine safeguards system performance. The data and algorithm are discussed in Section 2.

One step in the algorithm is to aggregate component performance to an intermediate level called "subsystem" performance. Subsystems are collections of components that have similar performance characteristics. For instance, in the current version of the ASM we aggregate components into three subsystems:

- Electronic detection
- Visual detection
- Accounting (records detection).

The aggregation scheme is shown in Fig. 2.

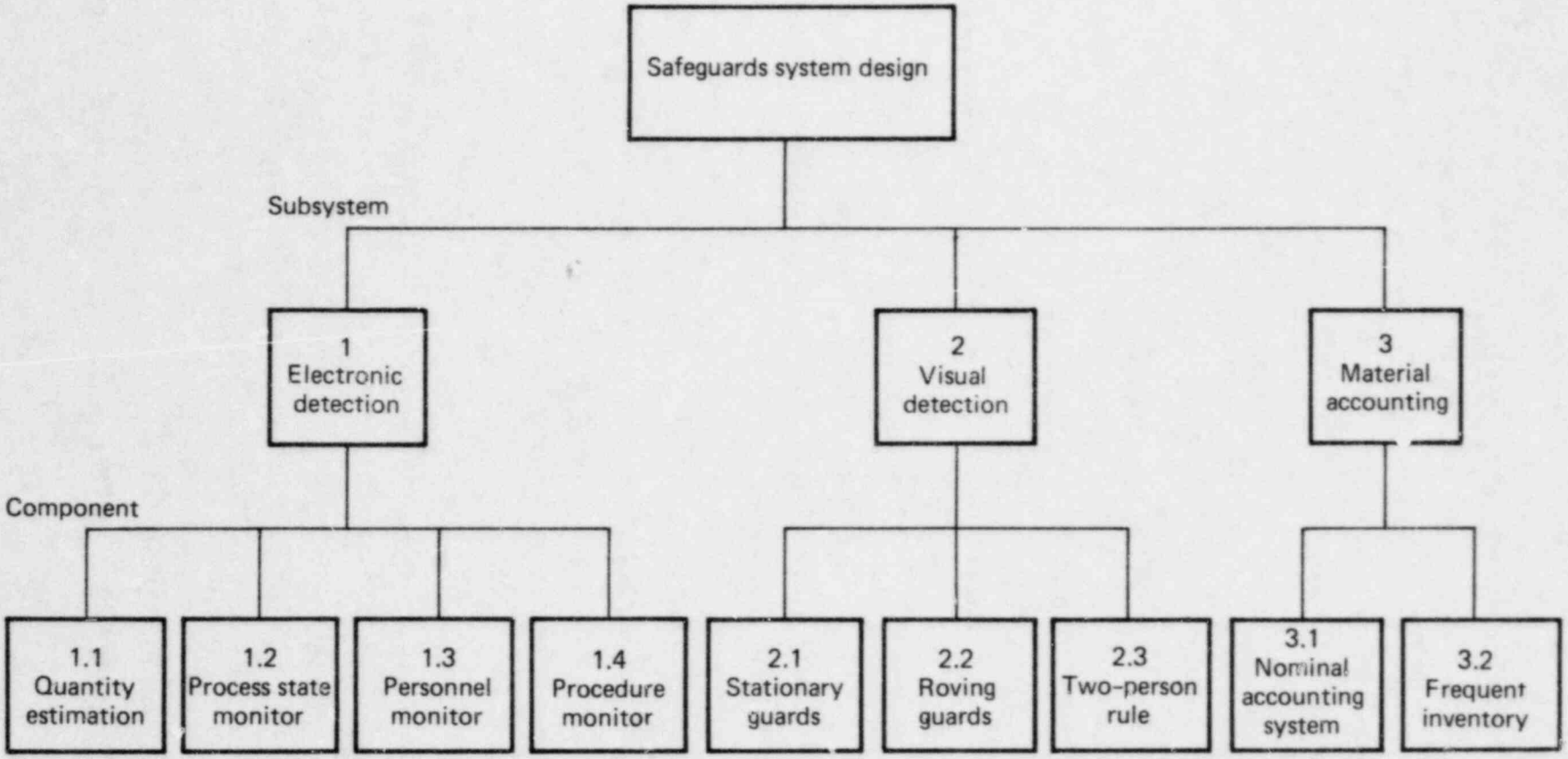


FIG. 2. Elements of a safeguards system design.

There are two reasons for the aggregation, which we explain below. Our reasons are based on a need to condition system identification probabilities and adversary abort decisions on the detection event. In other words, we believe that the probability of identification might be different if a guard detected the adversary than if the diversion is detected by a process monitor. Moreover, the adversary might behave differently in the two cases. This difference motivates the conditioning. But because we also believe that many detection events lead to similar abort decisions or identification probabilities, we condition the identification probabilities on aggregated subsystems rather than on individual components. The requirement that this conditioning places on the aggregation scheme is that all components in a subsystem must lead to the same identification probabilities or adversary abort/no abort decisions.

The first reason for aggregation relates to the safeguards authority's ability to identify the adversary and to confirm the attempt when a detection has occurred. The probability of this event is conditioned only on "which subsystem detected the adversary," not on "which component detected." Note that this implies that all components in a subsystem must lead to the same chances of identification. In other words, if an expert were assigning the probability that a given adversary, once detected, could be identified, the expert would need to know only whether or not one or more components in the subsystem made the detection; the expert would not need to know which component in the subsystem detected the adversary. If we relaxed this constraint in a model with nine components and three subsystems, we would have to condition the identification probabilities on the detecting component (with 512 combinations for each adversary) rather than on the detecting subsystem (with eight combinations for each adversary).

The second reason for aggregating components into subsystems is similar to the first. Recall that the current model allows the adversary to make strategic and tactical decisions. (These will be discussed in more detail in Section 2.) One tactical decision includes aborting the try if it has been detected by the safeguards. We condition this tactical decision on which subsystem detected the adversary, not on which component made the detection.

This places a second requirement on the aggregated components: detection by one or more components in a subsystem must lead to the same abort/no abort decision by the adversary. Relaxing this assumption again leads to 512 assessments instead of eight.

Since we assume that identification probabilities depend on the adversary's abort/no abort decision, the number of identification assessments is actually double the numbers given earlier. Thus, there would be 1,024 assessments for each adversary if the identification were conditioned on components and 16 if it were conditioned on subsystems.

ADVERSARY MODEL

Here we describe the types of adversaries currently in the model, the decisions they make about how to divert SNM, and the utility they assign to possible outcomes.

Adversary Types

The analysis began with 14 adversary types, whose characteristics are listed in Table 1. (See Ref. 6.) Sensitivity analysis showed that (in the illustrative data set) six of these adversaries dominated the other eight in terms of their contribution to the expected quantity of SNM diverted. Therefore, only adversaries 1, 3, 4, 9, 11, and 12 were retained for this analysis. These are identified with the superscript "a" in Table 1.

For all adversaries, we define an attempt as the existence of an adversary who desires to divert SNM. The attempt may consist of one or multiple tries, or, if the adversary is deterred, no tries at all.

Diversion Strategy and Tactics

We explicitly model the adversary's choice of diversion strategy (pathway to the SNM, number of tries, and the quantity to steal on each try) and tactics (abort the attempt or abort a try if detected). We assume that the

TABLE 1. Adversary characteristics.

Number	Access	Resources		Collusion	Desired quantity	Tries	Attempt frequency (att./1,000 y)	Percent of SNM diverted
		Equipment	Authority					
1[a]	Outsider	Major	No	No	Bomb	One	2.0	38.0
2	Outsider	Minor	No	No	Bomb	One	0.02	0.003
3[a]	Insider	Major	No	Yes	Bomb	One	0.29	4.7
4[a]	Insider	Major	No	Yes	Bomb	Multiple	2.6	33.0
5	Insider	Major	No	Yes	Less	One	0.03	0.049
6	Insider	Major	No	No	Bomb	One	0.03	1.1
7	Insider	Major	No	No	Bomb	Multiple	0.26	1.2
8	Insider	Major	No	No	Less	One	0.03	0.006
9[a]	Insider	Minor	No	Yes	Less	Multiple	13.0	13.0
10	Insider	Minor	No	No	Less	Multiple	1.4	0.14
11[a]	Insider	Major	Yes	Yes	Bomb	One	0.03	1.2
12[a]	Insider	Major	Yes	Yes	Bomb	Multiple	0.15	7.0
13	Insider	Major	Yes	Yes	Less	One	0.07	0.11
14	Insider	Major	Yes	Yes	Less	Multiple	0.11	0.18

[a]These adversary types are dominant.

adversaries make these choices based on their own preferences for outcomes such as capture, partial success, complete success, or failure without capture (assumed to be the same as no attempt at all).

The diversion strategy characterizes the kind of attempt the adversary makes, whereas the tactic specifies decision rules for aborting. The most important element in the strategy is the diversion path, also called a Monitor Target Set (MTS). These terms refer to the set of safeguards components that could feasibly detect the adversary enroute to the target SNM, while acquiring the SNM, or while leaving the facility with it. The diversion path or MTS will always be a subset of the components in the system design.

An important aspect of the adversary's strategic decision is whether or not to abort the entire attempt. With some diversion paths (strategies), there may be so many opportunities for detection that the adversary may be better off not trying. If this is true for all strategies given a "secure" system design, the adversary will be deterred, i.e., will make no attempt at all.

The adversary's tactical decisions are concerned with whether or not to abort the try if it has been detected. We assume that an insider adversary's chances of identification are lower if he or she aborts a detected attempt. In covert attempts, this means that they may be better off to abort. The tactics chosen by the adversary include a set of decision rules such as "abort if detected by Subsystem 2 but not if detected by Subsystems 1 and 3." Another might be "abort only if detected by both Subsystems 1 and 2." These tactics do not specify that every try will be aborted, but rather the conditions under which an abort decision will be in the best interest of the adversary. If those conditions occur, then the try will be aborted.

We assume that the adversary evaluates all possible strategies and tactics, and then chooses the combination that best serves his or her interest. These strategies and tactics are preselected; that is, the decisions are made before the first try; we assume that the adversary does not change strategy or tactics for the duration of the attempt.

Adversary Utility Model

We postulate that the adversary's choice among strategies and tactics is made based on maximum expected utility. Three factors matter to the adversary:

- Capture
- The status quo (no diversion)
- The quantity of SNM diverted.

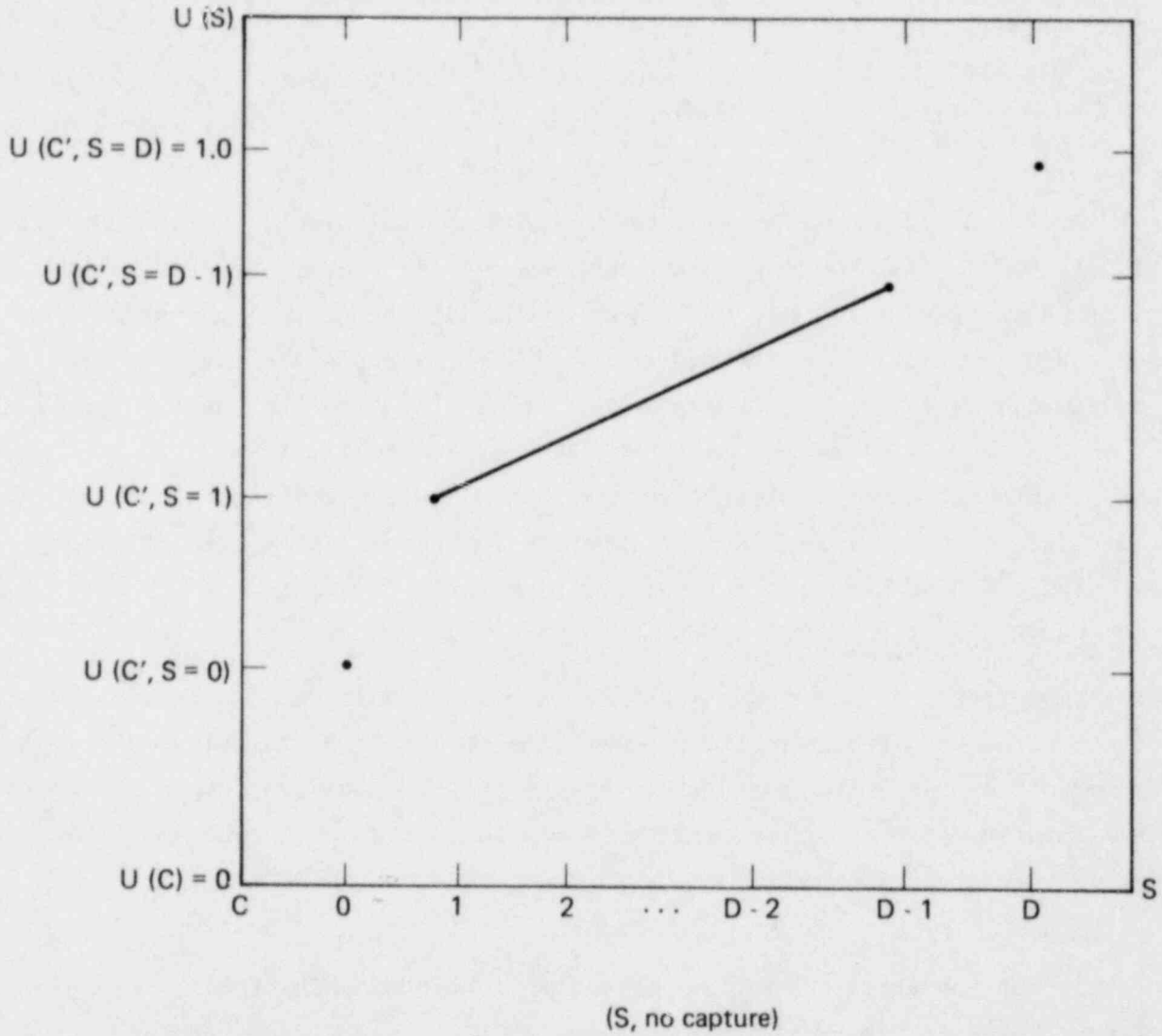
These three factors form the basis of the utility assessment. For an adversary trying to divert the SNM in one lump sum, we need be concerned only with the utility of capture, U(C), the utility of acquiring the desired amount D, U(S=D), and the utility of the status quo, U(S=0) where S denotes the number of successful diversions. The status quo represents the case where no try is made or the case in which no SNM is acquired before identification occurs. An adversary who repeats the attempt to steal many small quantities also has utility for partial success, U(S=m), where m represents the number of small quantities obtained before the adversary is identified.

Further, we assume that for the multiple quantity case, the adversary's utility function for diverted SNM is linear, with the exception of the first and last increments. The last increment might have more value than any other increment if the whole quantity is necessary for some goal. Figure 3 shows the form of a typical utility function.

Assessment of the utility function is accomplished using the three lottery questions shown in Fig. 4. Three numbers, P_1, P_2, and P_3, are required for each adversary. Given these three numbers, the utility function is:

$$U(C) = 0.0$$

$$U(S) = \begin{cases} P_1 & S=0 \\ aS + b, & 1 \leq S \leq D-1;\ D \geq 1 \\ 1.0, & S=D \end{cases}$$



C = Capture state
C' = No capture
S = Number of successful diversions
D = Desired number of successful diversions (D > 1)

FIG. 3. Adversary utility function.

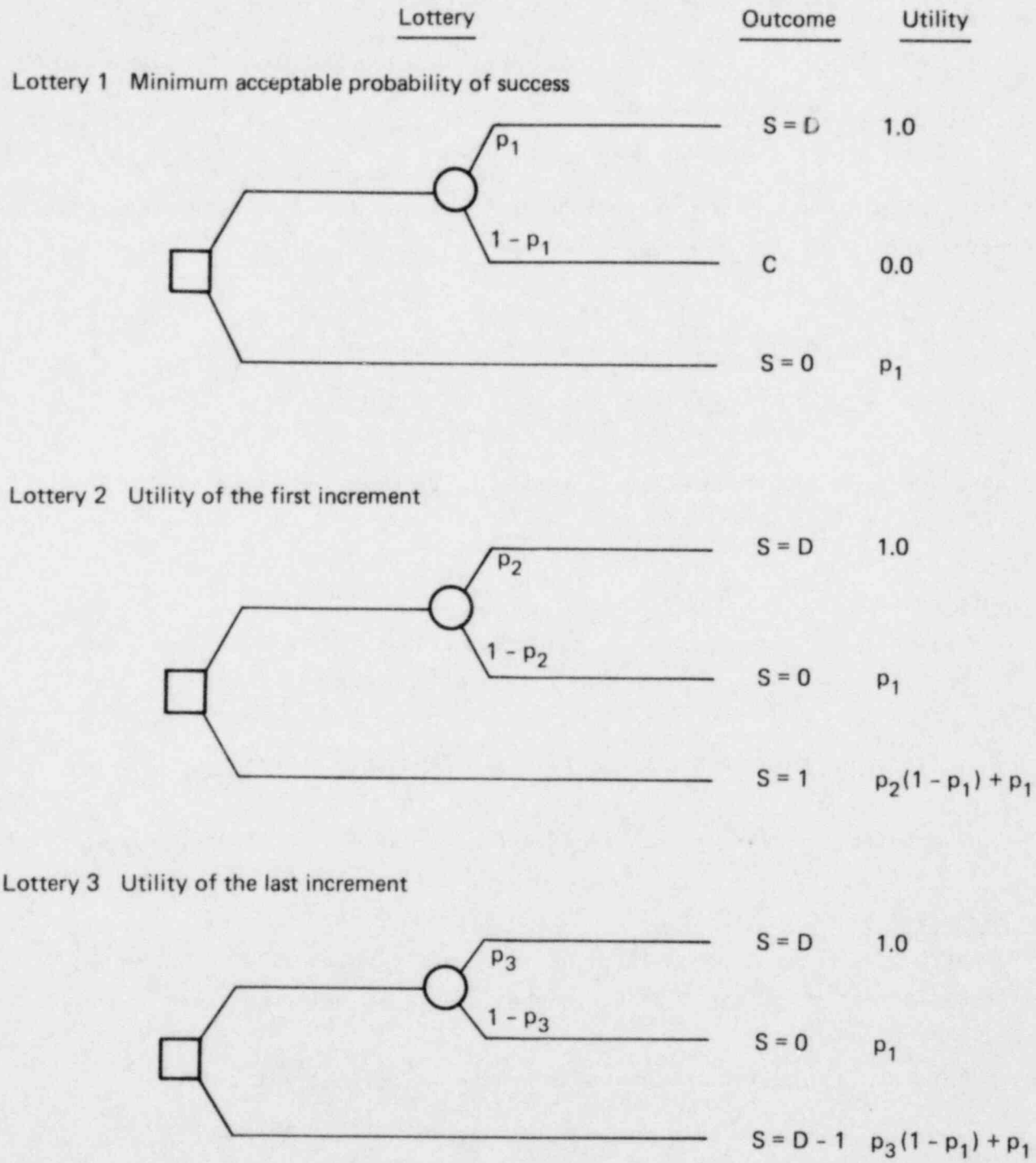


FIG. 4. Assessment lotteries for the adversary utility model.

where

$$a = \begin{cases} \dfrac{(1-P_1)\ (P_3-P_2)}{D-2} & , D>2 \\ 0 & , D=2 \end{cases}$$

$$b = \left\{ P_2(1-P_1) + P_1 - a \right. .$$

The coefficients a and b are derived by the straight line between points U(S=1) and U(S=D-1):

S	U(S)
1	$P_2(1-P_1) + P_1$
D-1	$P_3(1-P_1) + P_1$

The assessment questions posed by lotteries 1, 2, and 3 are easy to understand.

For Lottery 1:

Given only two outcomes, total success and capture, what is the minimum probability of success needed to induce an adversary's attempt?

A high value of P_1 means a strong aversion to capture.

Lotteries 2 and 3 apply only to adversaries making repeated attempts.

For Lottery 2:

What is the minimum probability of success for which an adversary would take an "all or nothing" gamble rather than keep only the first increment?

A high value of P_2 implies a high value of the first increment.

For Lottery 3:

What is the minimum probability of success for which an adversary would take an "all or nothing" gamble rather than keep all but the last increment?

A high value of P_3 implies a low value of the last increment. Also, we require $P_3 \geq P_2$. If not, then

$$U(S>1) < U(S=1),$$

or the adversary is worse off with more SNM.

Capture need not be the worst outcome from the adversary's perspective, as it is in Fig. 3. Assume for the moment that:

$$U(S=D) > U(C) > U(S<D),$$

where C, S, and D are as defined in Fig. 3.

Rational behavior for this adversary dictates selecting a strategy with maximum probability p(S=D) and minimum probability p(S<D). It is likely that such a strategy would be an "all or nothing" attempt, with D=1. This adversary will choose a strategy with maximum probability of success. Since this decision criterion is examined as a sensitivity case in Section 3, we will assume that all adversaries have the general utility function shown in Fig. 3. That is:

$$U(S=D) > U(0 \leq S<D) > U(C).$$

The linear utility function from U(S=1) to U(S=D-1) implies that the adversary is risk neutral for quantities in this range. In other words, the adversary is willing to "play the averages" for incremental gains in the region. Depending on the adversary's risk preference, this may understate or overstate the utility of some strategies for the adversary. However, given the flexiblity of utility assignments for U(S=0) and U(S=D), and the uncertainty as to the actual utility functions of adversaries, we believe the linear segment of the function is not a serious model limitation.

ADVERSARY-FACILITY INTERACTION

The interaction between the adversary and the facility safeguards is complex; therefore, we have modeled the interaction in some detail. Two factors make the interaction complex. First, the adversary will optimize his or her strategy to find the weakness in the safeguards system. As new safeguards are installed to foreclose one diversion path, we assume the adversary will shift to the next weakest link. Second, an adversary may make repeated covert tries, perhaps waiting until the system is vulnerable because of safeguards component failures. A safeguards component that is "slow but sure to detect" may be valuable in preventing diversion by a repeating adversary.

This section describes first how we model the adversary's choice of strategy. We then discuss a dynamic probabilistic model of repeated adversary trials.

Schematic Decision Tree Description

Figure 5 is a schematic decision tree showing major decisions and outcomes in the adversary-facility interaction. Square nodes represent decisions: those with an "S" in the box are social decisions (perhaps made by the facility designer following NRC guidelines) and those with an "A" are adversary decisions. Circles represent probability nodes, where the decision maker is uncertain as to which outcomes will occur. We show only two branches at each node, although there can be many.

The first decision we consider is the facility design--which safeguards components to choose. This decision is made without perfect foresight as to which adversaries (if any) will attack and whether the system can detect, identify, or interrupt the attempt.

Each adversary chooses a strategy and a tactic, as discussed earlier. This decision is the third node in Fig. 5. These decisions consider the probabilities of detection, identification, and interruption leading to the outcome states at the right-hand side of the figure. For some adversaries, the detection/identification/interruption sequence of events is repeated for several tries.

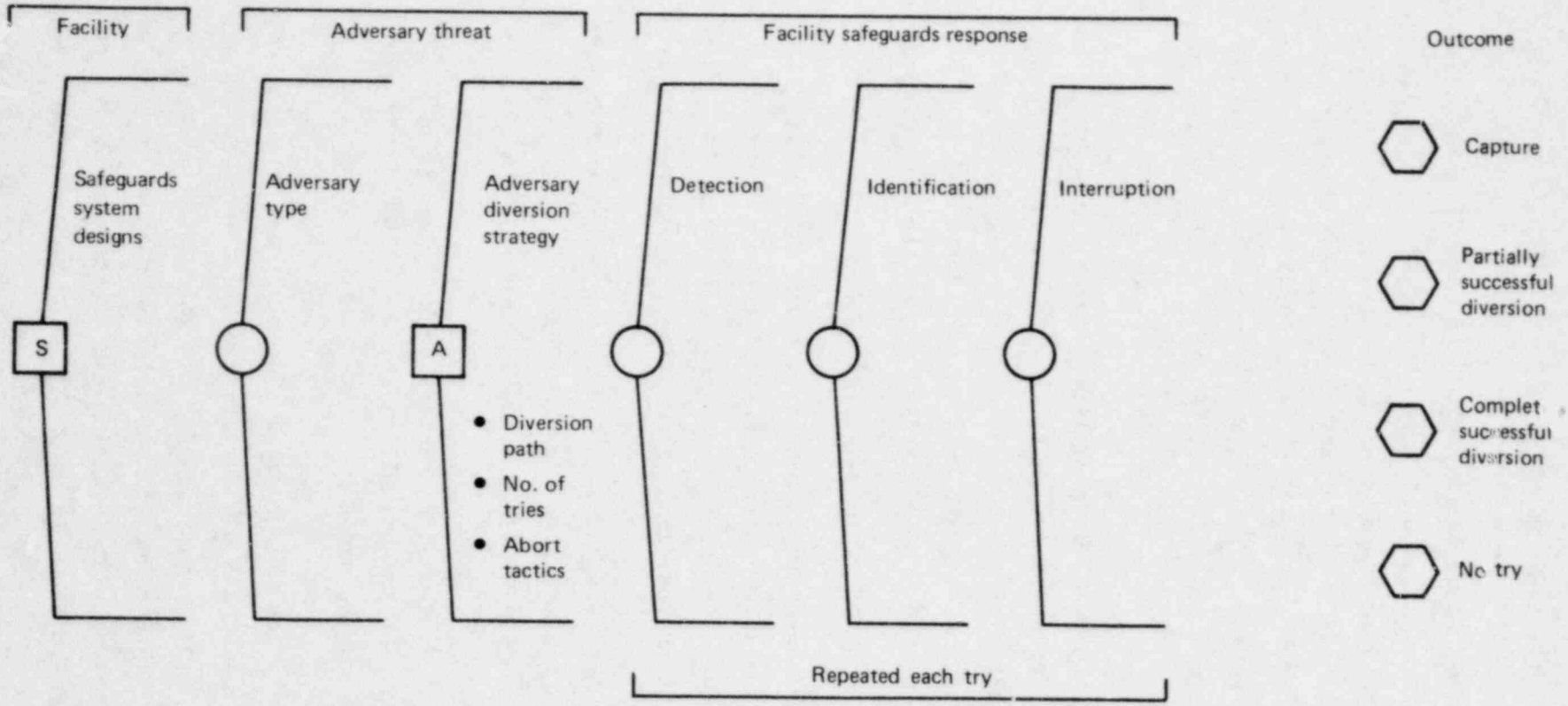


FIG. 5. Decision tree (adversary-facility interaction).

As defined in the Glossary, detection implies the transmission of an "attempt" signal to the safeguards authority. However, at that point the signal is not distinguished from a false alarm. "Identification" means that the authority knows the alarm is real and who the adversary is, so that the guard force can be dispatched to confront the adversary. "Interruption" means that no SNM crosses the plant boundary on a given try and the adversaries are captured. Because of this capture assumption, even if SNM were diverted on a previous try, no harm can come from it. If the adversary is identified but not captured, the attempt stops but the previously diverted SNM is still potentially lethal.

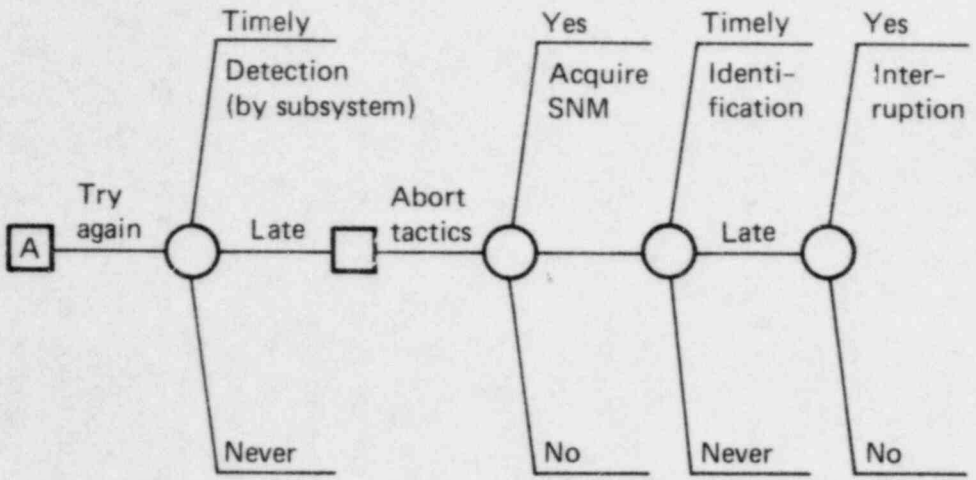
When using a dynamic model with explicit repeated attempts, identification is a crucial event. In order for physical security to interrupt the sequence, timely identification of the adversary sequence must occur. If the attempt is detected, aborted, but the adversary is not identified in time, then the attempt may be repeated. Nonidentification is the key to repeated attempts, and therefore it is included as an explicit event.

We assume that the adversary's perceived probabilities of detection, identification, and interruption are equal to those that would be assigned by the system designer. In other words, the designer and the adversary have the same state of information. This assumption could easily be relaxed.

We conclude this description of the schematic decision tree with a discussion of the sequence of events in a repeated attempt. These appear in Fig. 5 and are shown in greater detail (but still in schematic form) in Fig. 6.

The structure in Fig. 6 is repeated for each try. The nodes in the tree are all conditional on the adversary's chosen diversion strategy and tactics. We assume that these are constant for all attempts.

We assume that an adversary, whose method of diversion is by repeated attempts, will continue trying until all desired SNM has been diverted or until identification by the safeguards system. This is a result of the Adversary Utility Model, which assumes that utility is assigned only to possessing SNM; no "disutility" is associated with having to try again. For this reason, the adversary decision at the left side of Fig. 6 has only one branch: "Try again."



Outcome states	Adversary utility (j^{th} try)
Not repeatable	
Success (S_u)	$U(S_j)$
Failure (F_u)	$U(S_{j-1})$
Capture (C)	U(C)
Repeatable (limited)	
Success (S_{R1})	$U(S_j) + E[U(L)]$
Failure (F_{R1})	$U(S_{j-1}) + E[U(L)]$
Repeatable (unlimited)	
Success (S_{R2})	$U(S_j) + E[U(U)]$
Failure (F_{R2})	$U(S_{j-1}) + E[U(U)]$

S = Cumulative increments of diverted SNM in j tries (at least one try is successful)

E[U(L)] = Expected utility of a limited number of future attempts

E[U(U)] = Expected utility of an unlimited number of future attempts

FIG. 6. Repeated attempt sequence events (schematic tree).

We assume three forms of detection by the safeguards subsystems:

- Timely: with sufficient time to allow physical security to try to interrupt the sequences
- Late: eventual detection, but after the attempt is concluded
- Never.

The second adversary decision in Fig. 6, the tactical abort/no abort decision, is also shown with only one branch. This is to highlight our assumption that the adversary preselects a decision rule for abort/no abort before the first attempt, and sticks with the rule for all attempts. The rule specifies an "abort" or "no abort" decision for each possible detection outcome. An example of part of the rule is:

Detection by Subsystem:

1	2	3	Tactical Decision
Timely	Timely	Not timely	Abort
Timely	Not timely	Not timely	No abort
Not timely	Timely	Timely	Abort

If all three subsystems can give timely detection, then there are eight decision rules to be chosen before the first attempt is made. This set of eight decision rules is a tactic. Appendix A describes the tactical decision in more detail.

The next node in Fig. 6 is the binary "acquire SNM" probability node. The adversary either acquires the desired increment of SNM or not. We assume that the quantity of SNM sought on each attempt is the same. If the adversary aborts, no SNM is acquired.

Identification of the adversary by the safeguards authority can come timely, late, or never. If interruption is to occur, both detection and identification must be timely.

Seven outcomes states are possible with every attempt. They are classified as "not repeatable," "repeatable for a limited time," and "repeatable for an unlimited time." Within these three categories, there are at least two states representing success or failure to acquire SNM during the just-completed attempt. Table 2 shows which repeated attempt sequence events determine the

TABLE 2. Relation between outcome states and attempt sequence events.

Outcome states	Attempt sequence events
Not repeatable	
Capture	Interruption
Failure	Not acquire SNM, timely identification, not interrupted
Success	Acquire SNM, timely identification, not interrupted
Repeatable (limited)	
Success	Acquire SNM, late identification
Failure	Not aquire SNM, late identification
Repeatable (unlimited)	
Success	Acquire SNM, never identified
Failure	Not acquire SNM, never identified

outcome state. The identification event determines repeatability, and the "acquire SNM" event determines attempt success or failure. The "Success" and "Failure" states in the not-repeatable category are both instances in which the adversary is not captured, but because he or she is identified, the attempt cannot be repeated.

Finally, Fig. 6 lists adversary utilities assigned to each state. As explained above in Fig. 3, the adversary utility function has two dimensions:

capture and the quantity of diverted SNM. The adversary's utility at the completion of the j^{th} attempt is shown in Fig. 6, depending on which state is obtained. For nonrepeatable states, the utility is either U(capture) or U(cumulative SNM diverted). The utility of repeatable states is U(cumulative SNM diverted), plus the expected utility of future attempts. The maximum utility is U(desired quantity of SNM).

Extensive Form of the Decision Tree

Figures 5 and 6 showed the general form of events and decisions in the adversary-facility interaction. We now show explicitly the tree structure implied by the schematic decision tree. Using this extensive form of the tree, we will show how the dynamic probabilistic model is constructed.

To facilitate our description, we assume that the adversary has preselected a diversion strategy. Also, we assume that the decision rules (tactics) for aborting the try have been chosen. The spectrum of tactics ranges from "always abort," in which the try is aborted if any subsystem detects, to "never abort," in which the try continues regardless of which subsystem detects the try. For this example, we assume that the adversary's tactic is "always abort" if detected.

Figure 7 shows the decision tree that the adversary faces on each try. The adversary begins the diversion and is or is not detected by each of the three subsystems. We assume that the adversary knows if detection has occurred in a timely fashion: there is an audible alarm, flashing lights, the appearance of guards, etc. If the adversary is detected by a subsystem or a particular combination of subsystems, he or she responds with the action dictated by the preselected tactic.

If the adversary aborts the try, obviously no SNM is acquired. If the try is continued, the adversary is actually capable of acquiring SNM only part of the time. Depending on physical barriers, the adversary may be physically unable to acquire the SNM regardless of Material Control and Accounting (MC&A).

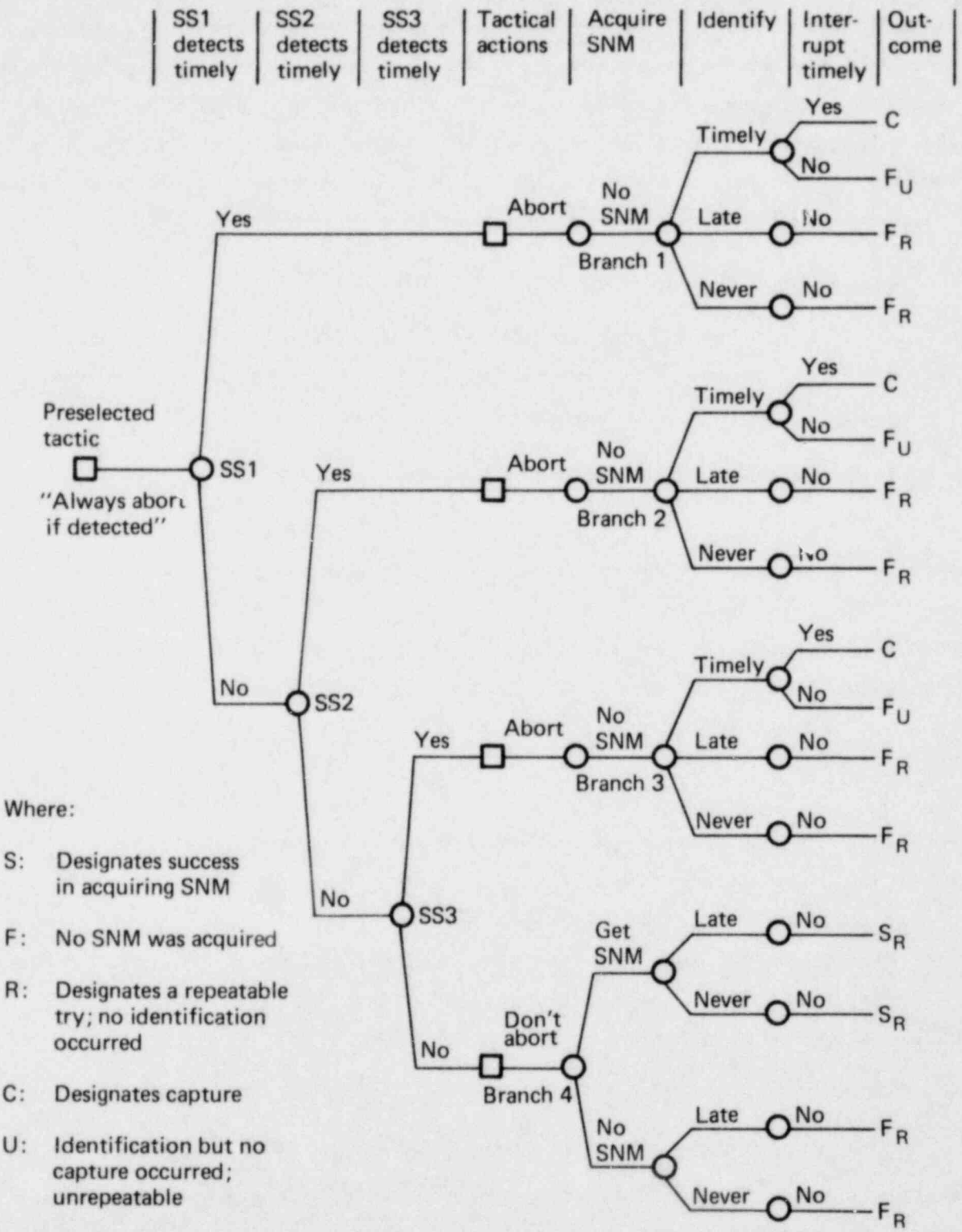


FIG. 7. Decision tree for each try.

The resulting probabilities of identification depend on which of the subsystems detected the adversary and whether or not the try were aborted. Logically, if there is no SNM in the adversary's possession, the probability that the safeguards authority can deduce that an attempt has been made will be lower. Perhaps the detection will be regarded as a false alarm. By always aborting if detected, an inside adversary who wishes to make numerous tries can minimize the probability of identification at the cost of not obtaining SNM on every try.

The next nodes represent identification and interruption. In order to interrupt the attempt and capture the adversary, the safeguards system must both detect and identify the adversary before leaving the facility.

The outcome states corresponding to the tactic "always abort" are listed on the right of Fig. 7. Notice that there are only a few branches that result in capture, and that the number of branches that yield any SNM is also small.

Figure 7 represents the decision tree for the tactic "always abort." For each strategy, there are various tactics that could be similarly depicted. Generating all possible tactics requires a lengthy explanation, which is given in Appendix A.

A further complication, not show in Fig. 7, is the probability that a subsystem detects an adversary at a later time, after the completion of the try. If the adversary is identified after the late detection, he or she will be apprehended at the next SNM diversion try. Appendix A also explains how this factor is included in the calculations of final success or capture probabilities for the adversary.

Markov Model

The tree in Fig. 7 can be coalesced into a lottery over the outcome states shown in the right-hand column. Consider the simple case in which the probability of late detection equals zero. Five outcome states result: capture (C), unrepeatable success or failure (S_U, F_U), and repeatable

success or failure (S_R, F_R). The probability of ending up in each of these states is the sum of the probabilities of each branch in Fig. 7 that corresponds to that state. The first probability node in Fig. 8 represents the coalesced lottery from Fig. 7.

If the adversary is able to repeat the try, he or she faces exactly the same lottery again. The crucial assumption is that probabilities do not change with each repetition. Of course, if the adversary is captured or identified, the process stops. At the end of each repeatable branch, the same lottery could be attached. The second and third nodes in Fig. 8 demonstrate how the decision tree could be repeatedly expanded. This quickly results in a large and intractable tree. Rather than attempting to solve this tree, we modeled the repeating adversary's iterative and probabilistic process as an equivalent Markov process.[3] Figure 9 presents this model; note that the Markov states and probabilities correspond directly to the states and probabilities in the outcome lotteries in Fig. 8. Also note that three states, F_U, S_U, and C are trapping states at which the process stops.

Markov processes are useful for modeling not only because of their explicit consideration of both the probabilities and the dynamic aspects of the case, but also for their ease in solving to achieve various results. Using the mathematical formulas associated with Markov processes, we can find the long-run probabilities of ending up in each of the outcome states. This calculation can produce a probability distribution over the final outcomes of the repeated trial. For example, suppose that an adversary intended to steal some number (D) of small increments of SNM (S), and that the adversary tried until the D amounts are acquired, the adversary is either captured, or is identified and unable to repeat. In any of these cases, the try would end. The Markov model would yield the probability of being in each of these stopping states, as well as the number of times the try was successful. The Markov model produces, therefore, a new lottery, which represents all final outcomes of the repeated process in Figs. 8 and 9. This is labeled "Outcome Lottery" in Fig. 10. Compare Figs. 5 and 10. Note that the "Outcome Lottery" in Fig. 10 now replaces the three repeated nodes: detection, identification, and interruption in Fig. 5.

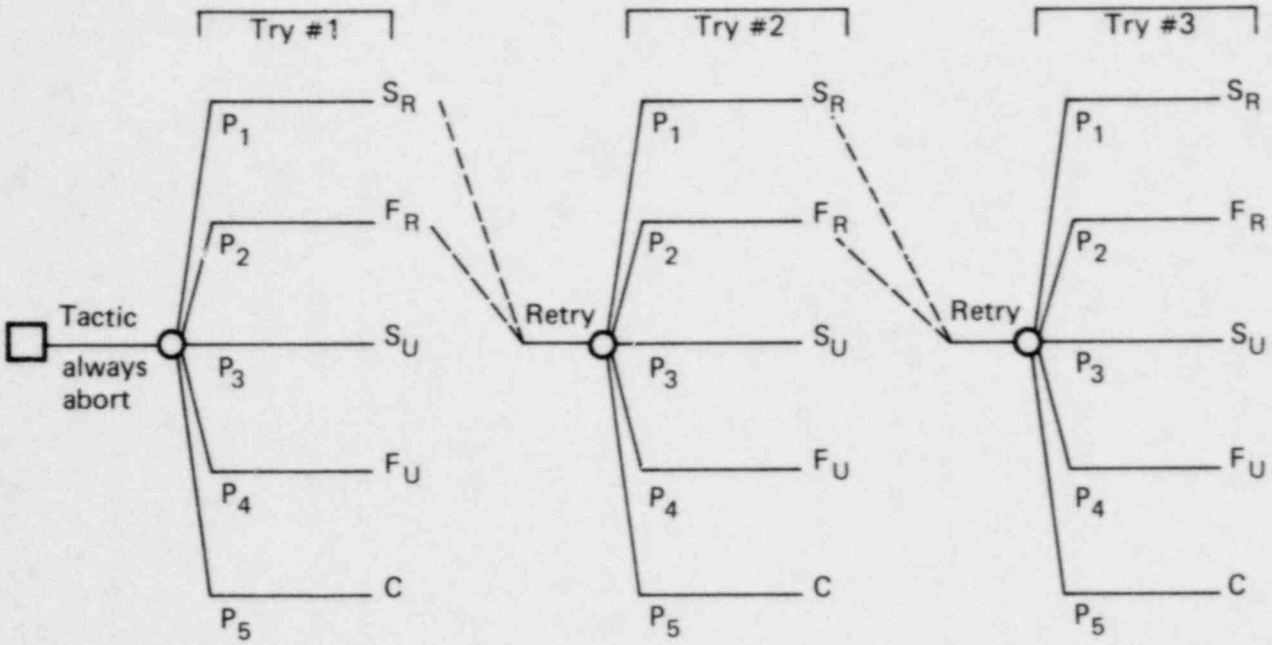


FIG. 8. Repeating lottery on the outcome states from each try.

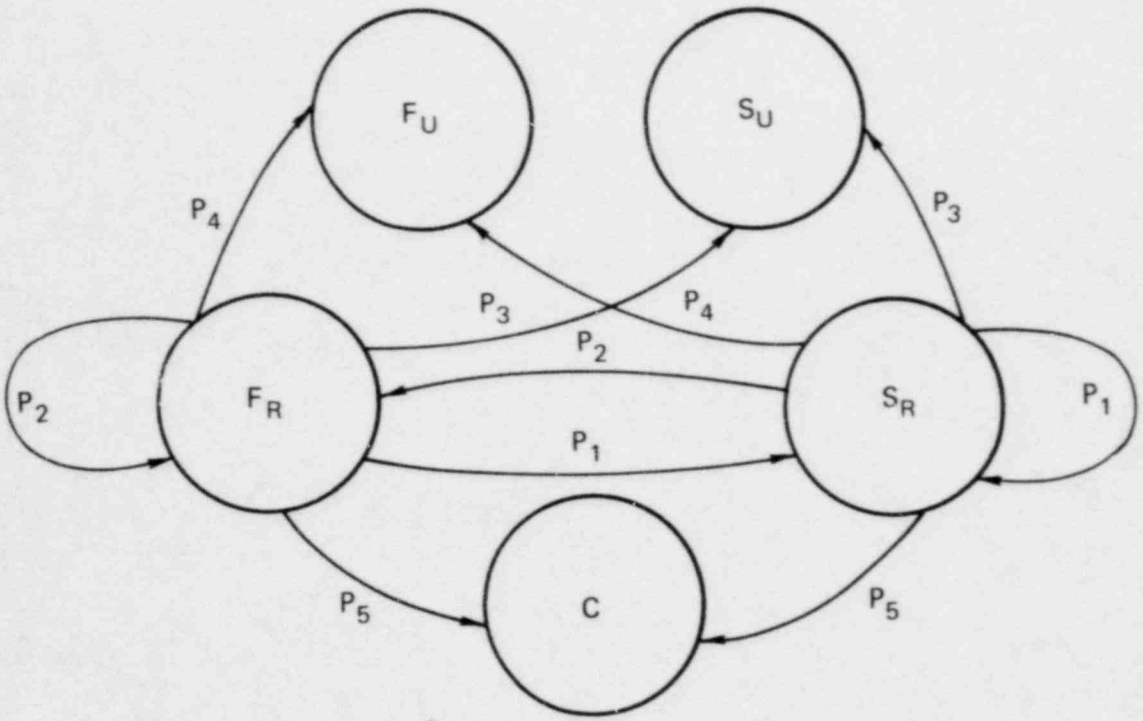


FIG. 9. Markov model.

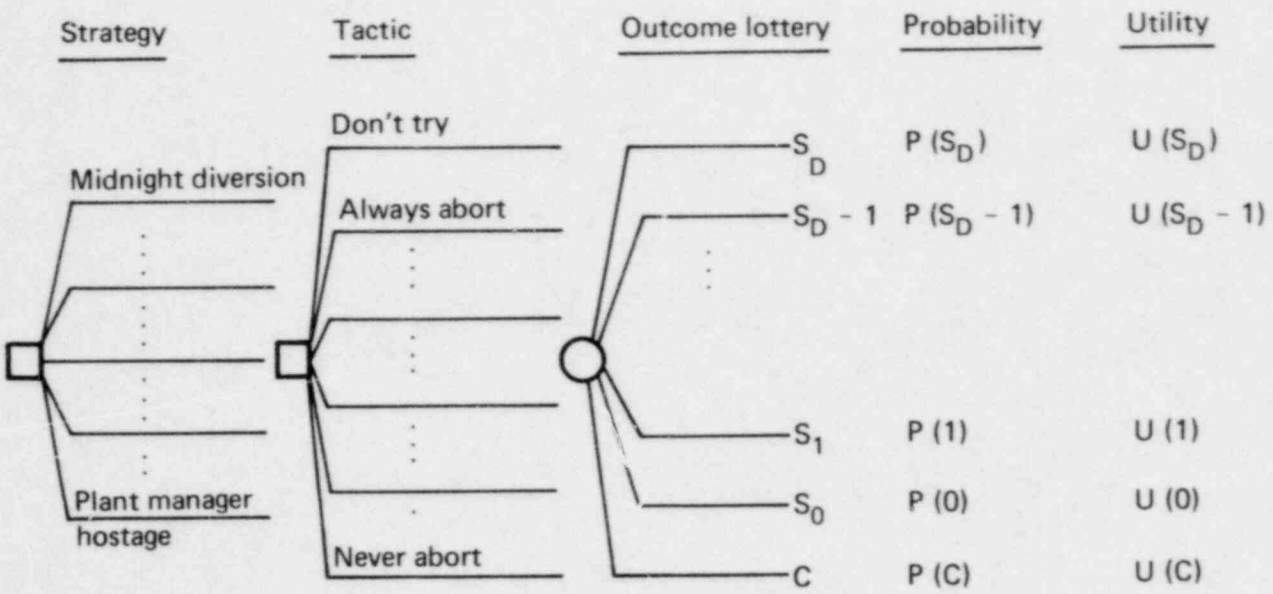


FIG. 10. Decision tree representing final outcomes from repeated trials.

Figure 11 contains the equations used to calculate the long-run probabilities for Fig. 10. Section I on Fig. 11 gives the probability of stealing D increments and the probability of being captured. The probability of being only partially successful--P(F)--is the sum of all the probabilities of stealing amounts less than D but being identified before D units were acquired. The distribution on the number of units of SNM diverted before identification occurs is found in Section II of the same figure, corresponding to the probabilities of S=D-1,S=D-2, ... S=1, S=0.

Notice that in Fig. 10 each outcome state has an adversary utility associated with it. The adversary can calculate the expected utility by multiplying the probability of each outcome times the utility of each outcome and summarizing all the products. This corresponds to the equation in Fig. 11. For a given strategy, the adversary can compare the expected utility for any tactic with the utility of each of the other tactics, and choose the tactic with highest utility. This is the best that can be done for that strategy. Using this, the adversary can compare all the strategies, and find the strategy and its associated tactic that will yield the maximum utility. Figure 10 shows the structure of the adversary's decisions.

Figure 11 contains one more interesting statistic. Section IV lists the equation giving the expected number of tries the adversary will make before he or she is identified by the safeguards system (and the process stops).

Although we did not discuss it here, we must consider the possibility of late detection and identification. Appendix B contains an explanation of how we included it in the formulation of the ASM. The equations used to solve for the long-run probabilities in the outcome lottery become much more complicated, but conceptually they are similar to what was explained above.

Once it produces the optimal solution for the adversary, the model can be used to choose an optimal design, as shown in Fig. 5. In addition, the information can be used by an NRC regulator determining how well a facility meets licensing or performance criteria. How well a system performs against an adversary who knows the best way to divert SNM is the most stringent test.

Outcomes after infinite iterations

I. Long-run probabilities:

$$P(D) = \left(\frac{P_1}{1-P_2}\right)^D \cdot \frac{\left(P_5(P_1+P_2) + (P_3+P_4)\right)}{(P_1+P_2)\ (P_3+P_4+P_5)}$$

$$P(C) = \frac{P_5}{(P_3+P_4+P_5)} \left(1 - \left(\frac{P_1}{1-P_2}\right)^D\right)$$

$$P(F) = \frac{(P_3+P_4)}{(P_3+P_4+P_5)\ (P_1+P_2)} \left((P_1+P_2) - \left(\frac{P_1}{1-P_2}\right)^D\right)$$

II. Expected utility to adversary:

$$EU = U(C) \cdot P(C) + U(D) \cdot P(D) + \sum_{m=0}^{D-1} P(m \ \& \ F) \cdot U(m)$$

III. Distribution on SNM diverted:

$$P(m) = \frac{P_1^m \cdot \left(P_5(P_1+P_2) + (P_3+P_4)\right)}{(1-P_2)^{m+1}\ (P_1+P_2)\ (P_3+P_4+P_5)}$$

IV. Expected number of tries:

$$E(n) = \frac{1}{P_3+P_4+P_5}$$

Number of successful tries:

$$= \frac{P_1}{P_3+P_4+P_5}$$

Number of failures:

$$= \frac{P_2}{P_3+P_4+P_5}$$

FIG. 11. Equations for computing Markov statistics.

Another measure of performance is to assign social utility to each of the adversary's outcome states using the Consequence Model discussed in the next section. For any given system design, this model produces a probability distribution over consequences (such as deaths, damages, etc.). After assigning utilities to consequences, one can roll back the tree in Fig. 5 to compute an expected utility for any particular design. Obviously, society will rate the utility of adversary capture very positively and will consider all the negative possible consequences associated with diverted SNM. Repeating the process for different designs identifies the design with the highest expected utility.

The Markov model also calculates the values of various performance measures that can be used for facility evaluation. Assuming that the adversary will use optimal strategy and tactics, system responses such as the probabilities of detection, identification, and capture, as well as the expected SNM diverted, can be calculated.

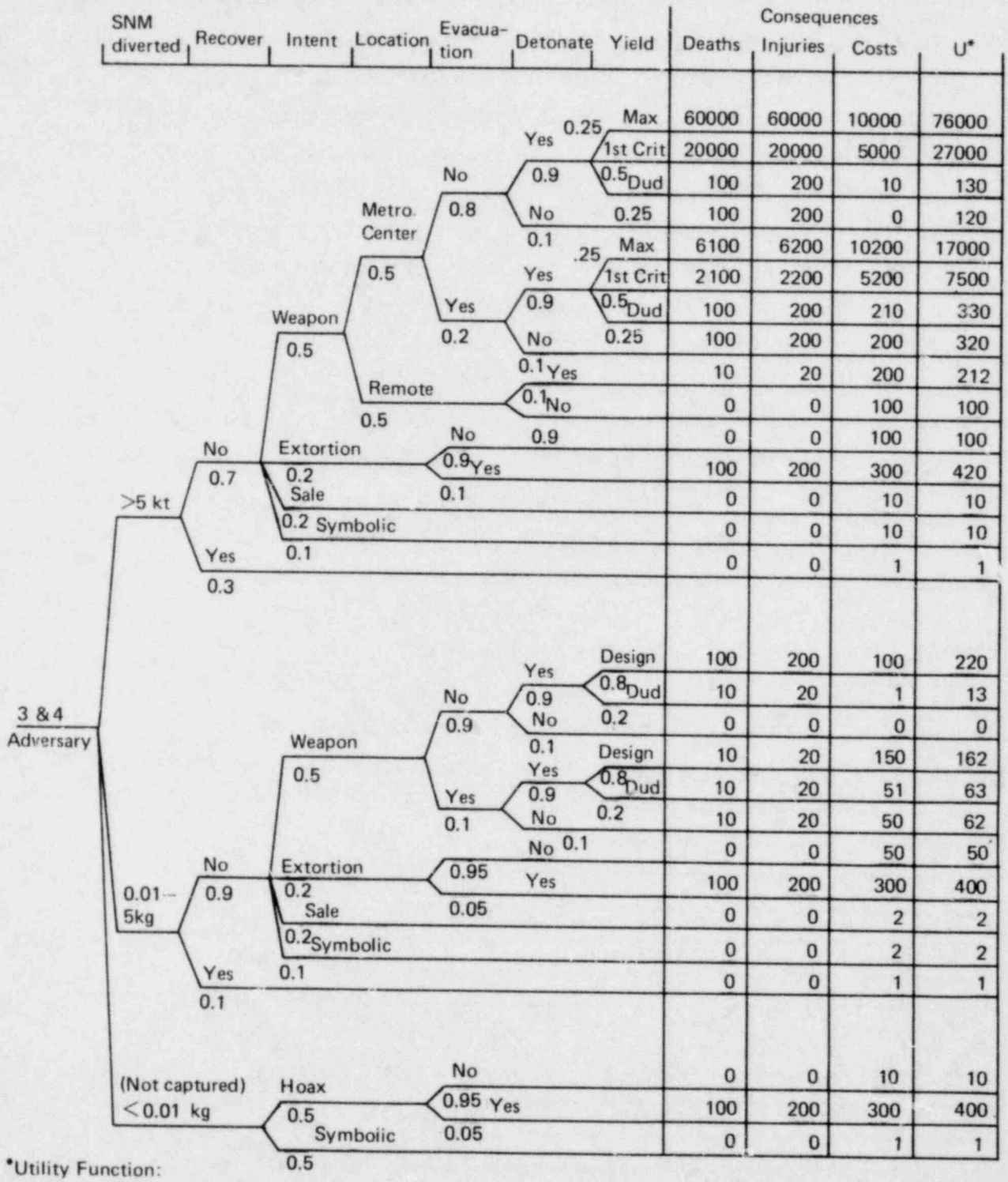
CONSEQUENCE MODEL AND SOCIAL UTILITY MODEL

Figure 12 shows a consequence probability tree with illustrative data. Six uncertain events are considered in the Consequence Model:

1. The intended use of the material
2. The success in making the nuclear device
3. The location of the resulting nuclear incident
4. Whether or not the local population is evacuated
5. Whether or not the device is detonated
6. The yield.

The following are some of the important features of this Consequence Model:

- Consequences are conditional on adversary.
- There is a probability that the SNM will be recovered before consequences occur.
- Weapon detonation consequences are a function of device yield.
- Hoax diversions followed by weapon threats are modeled.
- Injuries are postulated in addition to deaths.
- The "Intent" node includes symbolic diversions.[4]



*Utility Function:

1 X Deaths + 0.1 X Injuries + 1 X Costs = Utility (Equiv. 10^6)

FIG. 12. Illustrative consequence model.

We have conditioned consequences on adversary characteristics to reflect two dependencies on the adversary:

- Probability of recovery
- Probabilities over intents (weapon, extortion, sale, or symbolic diversion).

The probability of building a successful weapon may also be dependent on the adversary; however, given a successful 10-kg diversion, we assume that the adversary is sufficiently competent. This assumption could be changed easily. Table 3 shows the probabilities assigned to dependent events for the six adversary types. These assignments replace those in Fig. 12 where appropriate. Notice that the probability of recovery equals 1.0 for Adversary 9. This adversary is merely out to prove that SNM can be stolen. By definition, Adversary 9 is stealing at most 0.3 kg, and therefore will never have 5 kg, which is computationally equivalent to recovery with probability 1.0.

The consequences of weapon detonations are purely illustrative. However, a report soon will be issued by Dr. Dean Kaul at Science Applications, Inc. (SAI)[5] describing a model to predict the consequence of such events. Specifically, his model predicts immediate deaths and injuries due to blast, radiation, and thermal effects from a nuclear weapon detonation in two typical U.S. cities.

SAI's model could be used to derive 20 of the specific consequence numbers in Fig. 12. Using the model for sensitivity studies would allow conversion to probability distributions of the single point consequence estimates in Fig. 12. Because SAI's results are not yet published, we have not done any additional work on the consequences of these events. The assumption implicit in our probability assignments in Fig. 12 is that these bomb consequences are the dominant effects in the Consequence Model.

Table 4 shows the expected values of public and private consequences, as a function of quantity diverted. The public consequences are computed by rolling back the probability tree in Fig. 12 with probabilities assigned in Table 3. The right-hand columns give the three-point consequence utility function for each adversary and diversion quantity.

TABLE 3. Consequence model probabilities.

	Quantity diverted							
	5 kg				0.01-5 kg			
Adversary	P (Recover)	P (Weapon)	P (Extort.)	P (Sale)	P (Recover)	P (Weapon)	P (Extort.)	P (Sale)
1	0.5	0.75	0.10	0.10	0.40	0.75	0.10	0.10
3 and 4	0.3	0.50	0.20	0.20	0.10	0.50	0.20	0.20
9	1.0	-	-	-	0.90	0.10	0.50	0.10
11 and 12	0.3	0.10	0.10	0.50	0.10	0.10	0.10	0.50

TABLE 4. Expected consequences by adversary and quantity diverted (equivalent 10^6).

Adversary	Public consequences Qty. diverted (kg)			Plant consequences Qty. diverted (kg)			Total consequences Qty. diverted (kg)		
	<0.01	0.01-5	>5	<0.01	0.01-5	>5	<0.01	0.01-5	>5
1	$16P_e$	76	4696	13	13	20	$16P_e$ + 13	90	4700
3 and 4	$16P_e$	84	4396	14	14	10	$16P_e$ + 14	100	4400
9	$16P_e$	6	0	10	1	0	$16P_e$ + 10	7	0
11	$16P_e$	22	890	10	10	1	$16P_e$ + 10	32	890

P_e = 1-P (Capture)

II. EXAMPLE ANALYSIS

PURPOSE

This example analysis is designed to illustrate application of the methodology to setting safeguards criteria for a hypothetical nuclear fuel cycle facility. To make the example realistic, the numbers were developed by a few members of the LLL project team. The data reflect the subjective judgment of these individuals; they were not developed by detailed analysis, and they should not be regarded as accurate.

ASM INFLUENCE DIAGRAM--PROBABILISTIC DEPENDENCE

The ASM is a dynamic, probabilistic model, and as such has probability distributions for numerous variables in the model. Where two variables are dependent (for instance, the probability of detection and adversary type), we made conditional probability assignments. We use the influence diagram in Fig. 13 to show this dependence among variables.

Each square and circle in Fig. 13 represents a decision and an uncertain event, respectively. A single box denotes the regulator's or designer's decision; the double box represents the adversary's decision. Arrows between two circles or squares represent dependence in the model. Consider the identification event in the center of the diagram. The probabilities assigned to identification--"Timely, Late, or Never"--are conditional on the adversary, the adversary's tactic--abort or not--and whether detection is timely or not. However, the identification probabilities are independent of system design, the diversion strategy, and whether or not the adversary acquires SNM. Although in reality there may be some dependence on these other factors, we are making the assumption that all essential information for determining identification probabilities is contained in the detection, tactics, and adversary nodes.

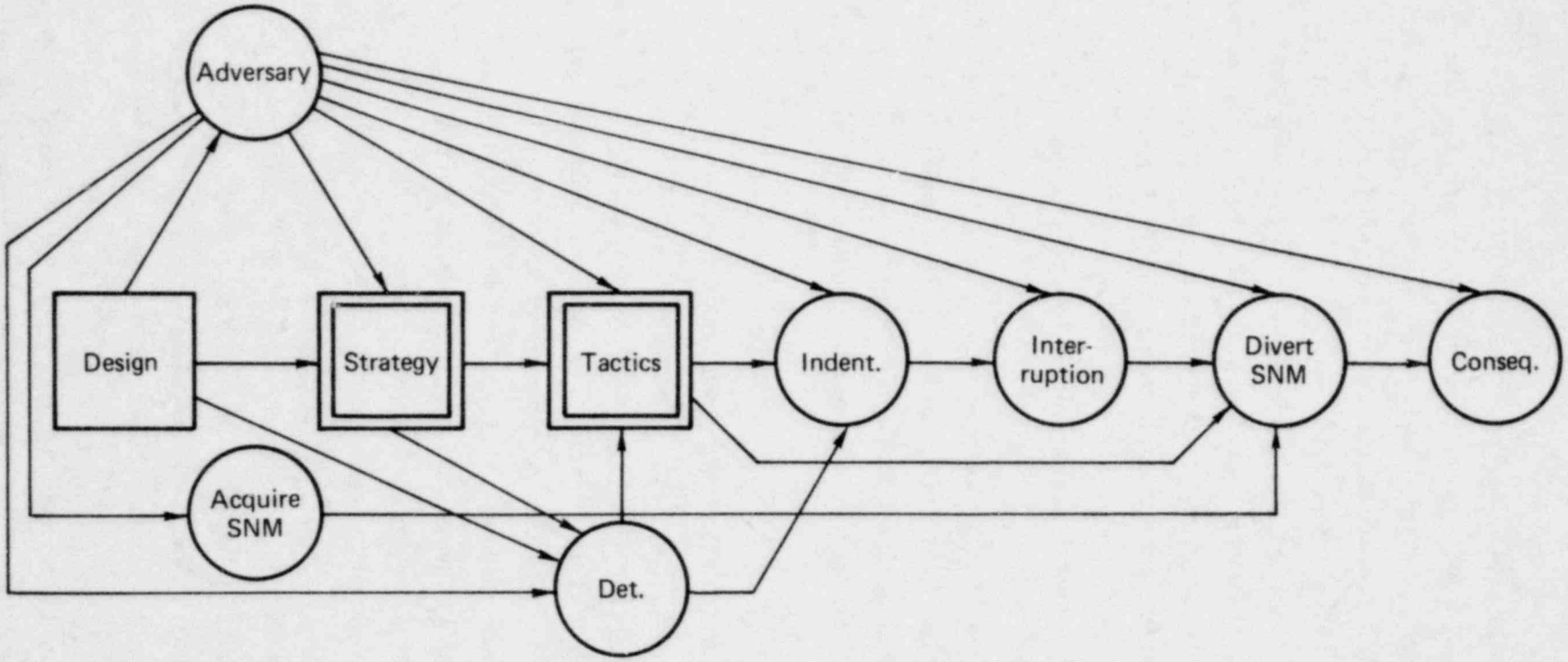


FIG. 13. Diversion model influence diagram.

Some important independence assumptions are:

- The adversary decides whether or not to abort, knowing that detection has occurred, but not knowing if he or she has been identified.
- Interruption probabilities depend only on identification, and not on how the adversary was detected.
- The acquisition of SNM depends on the adversary but not on the diversion strategy; that is, all feasible strategies will allow the adversary to acquire the SNM with the same probability.
- The adversary knows the plant design, at least to the extent to which safeguards components are included.
- The adversary may choose not to try if the design is particularly secure.
- The system design influences the probability of detection, but not the probabilities of identification or interruption; in other words, changing the design can improve the system's ability to detect an adversary, but cannot change the probability of identification or interruption.

DETECTION: AGGREGATING COMPONENT PERFORMANCE

Figure 14 shows additional details for the detection event in Fig. 13. The actions of an adversary are detected by safeguard components. In Fig. 14, Subsystem 1, SS1, is composed of Components C_{11}, C_{12}, C_{13}, and C_{14}. The probability that any component will detect the adversary is independent of detection by any other component (in the same or different subsystem). This is indicated by the lack of arrows among components in the large component performance circle. The probability of detection by any subsystem is the probability that one or more of its components detect the adversary. Figure 14 also shows that subsystem detection probabilities are independent of detection by components belonging to other subsystems. Thus, subsystems can have no components in common. This probabilistic independence among subsystems sets a constraint on the analytical process of aggregating components.

Figure 14 also highlights the assumption that adversary tactics and identification are dependent on subsystem, not component, detection. Making this independence assumption greatly simplifies the algorithm for computing

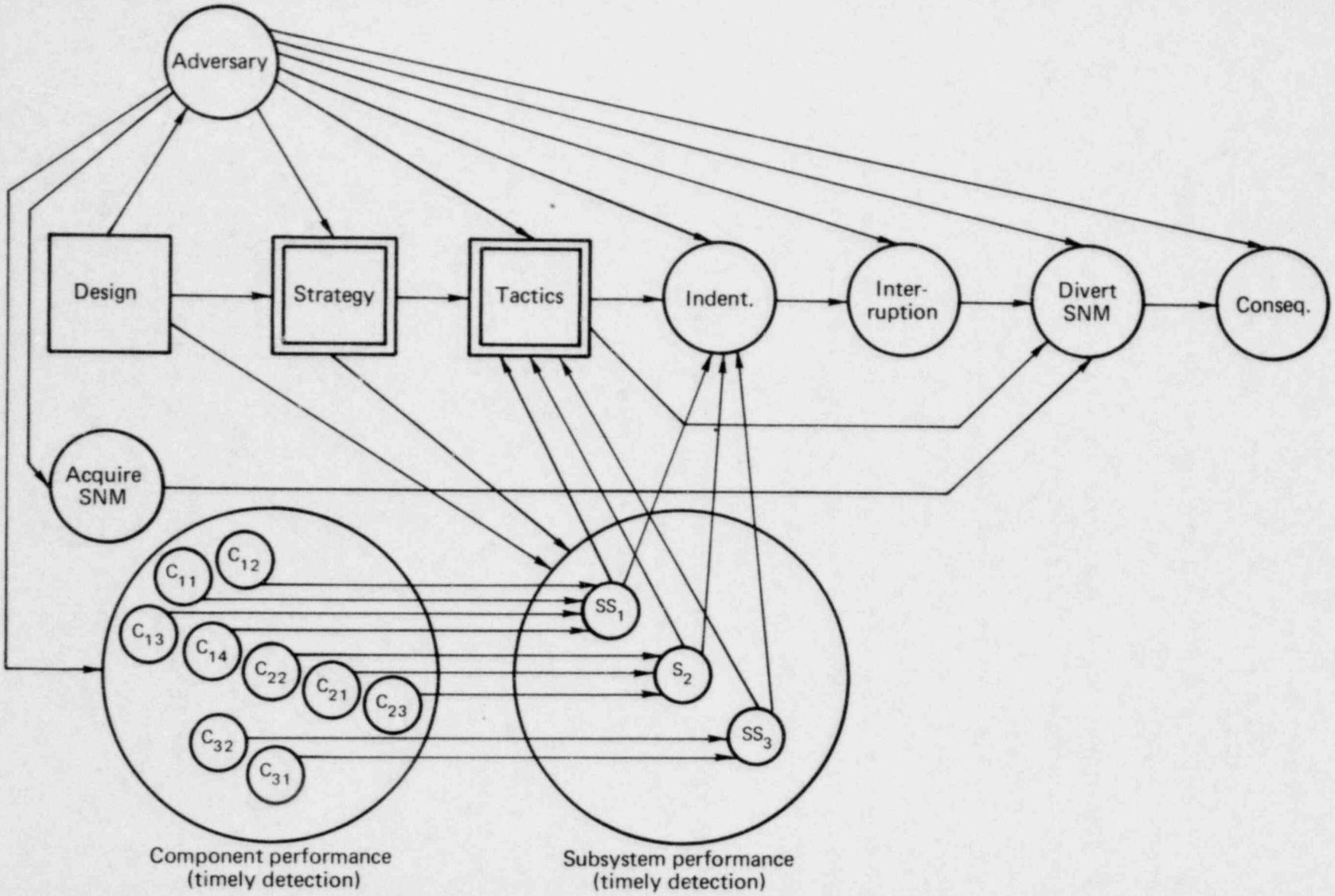


FIG. 14. Influence diagram with detection system detail.

the probability of detecting by a subsystem:

$$P(SS_i \text{ detects}) = 1.0 - P(SS_i \text{ not detect})$$
$$P(SS_i \text{ does not detect}) = \prod_j (1.0 - P(\text{Component}_{ij} \text{ detects})),$$

where i is the index for subsystems, and j is the index for components included in subsystem i.

Although it simplifies the calculation of subsystem detection probabilities, the independence assumption is not mandatory in the ASM. It could be relaxed in one of two ways:

- Direct specification of subsystem detection probabilities
- Specification of subsystem detection probabilities for all possible component configurations.

The first approach avoids altogether the use of components. Designs and adversary strategies are specified in terms of the subsystems included, rather than the components of subsystems.

The second approach requires computing subsystem detection probabilities external to the ASM. This would be done by an algorithm for handling the stochastic dependence among components. The external model would compute the probability of subsystem detection, where the subsystem contains only components that are both in the facility design and in the adversary's diversion path. This is done by considering every combination in every subset of components in the subsystem. For Subsystem 1 in the current model, which has four components, this requires 15 external calculations (assuming no late detection):

Number of calculations	Components in diversion path
4	C_{11}, C_{12}, one at a time
6	C_{11}, C_{12}, two at a time
4	C_{11}, C_{12}, three at a time
1	C_{11}, C_{12}, four at a time
15	

If late detection were possible, 30 probabilities must be specified for Subsystem 1: P_D (Timely), and P_D (Never) for all combinations of components. As long as the number of components is relatively small, this latter approach is not formidable.

Safeguards System Designs

The influence diagram in Fig. 14 shows a dependence between subsystem performance and the safeguards system design. As we have mentioned, system designs are combinations of components. For instance, three sample systems are:

System design	Components included
- Minimal design	C_{12}, C_{21}, C_{31}
- Maximal design	C_{11}, C_{12}, C_{13}, C_{21}, C_{22}, C_{31}, C_{32}
- Intermediate design	C_{12}, C_{13}, C_{21}, C_{22}, C_{31}

The system's design determines which components are included in the safeguards system, and, therefore, influences the subsystem probabilities of detection.

Table 5 is a matrix showing the example safeguards system designs. The nonzero entries in the matrix designate components included in the design (one row of the matrix for each design). The numbers 1, 2, and 3 indicate subsystem number. Designs 1 through 4 make up the base case and span the space of feasible designs. The remaining designs in Table 5 are used for sensitivity analysis.

Assumed component costs for a high throughput facility are shown at the bottom of Table 5. They are listed in thousands of dollars per year.

Adversaries

Adversary descriptions and illustrative attempt frequencies were shown above in Table 1. Table 6 below gives the probability that each adversary will acquire SNM.

Diversion Strategies

A diversion strategy includes a target quantity of SNM per attempt, an attempt frequency, and a diversion path or Monitor Target Set. Recall that the diversion path is the collection of safeguards components that could detect the

TABLE 5. Safeguards system designs.

Subsystem:	SS1--Material control				SS2--Physical security			SS3--Accounting			
Component:	1.1	1.2	1.3	1.4	2.1	2.2	2.3	3.1	3.2	Inspection per year	Description
Designs:											
1	1	1	1	1	2	0	2	3	0	12	Testbed design
2	0	0	0	0	2	0	0	3	0	6	Minimal
3	1	1	1	1	2	2	2	3	3	52	Maximal
4	0	1	1	0	2	0	0	3	0	6	Moderate
5	0	1	1	0	2	0	2	3	0	6	Physical security (P.S.) - A
6	0	1	1	0	2	2	0	3	0	6	P.S. - B
7	0	1	1	0	2	2	2	3	0	6	P.S. - C
8	0	1	0	0	2	0	0	3	0	6	Material control (M.C.) - A
9	1	1	0	0	2	0	0	3	0	6	M.C. - B
10	0	1	1	0	2	0	0	3	0	6	M.C. - C
11	1	1	1	0	2	0	0	3	0	6	M.C. - D
12	1	1	1	1	2	0	0	3	0	6	M.C. - E
13	0	1	1	0	2	0	0	0	0	6	Accounting - A
14	0	1	1	0	2	0	0	3	3	6	Accounting - B
Cost 10^3 per year:	2,450	600	600	2,000	300	400	500	300	5,000		

TABLE 6. Adversary probabilities for acquiring SNM.

Adversary Number	Description	P (acquire SNM/attempt)
1	Outsiders	0.7
3	Insider; major equipment; one attempt	0.8
4	Insider; major equipment; multiple attempts	0.8
9	Insider; minor equipment; less than bomb quantity	0.1
11	Insider; major personnel; one attempt	0.8
12	Insider; major personnel; multiple attempts	0.8

P(Attempt) = 0.02

adversary during the attempt. If, for example, the target quantity is less than the detection threshold of the component, then the component is not in the diversion path. Table 7 lists the diversion strategies for each adversary; a nonzero entry in any row means the component is included in the path. Only adversaries 4, 9, and 12 use multiple attempts. For these perpetrators, Table 7 shows diversion strategies with varying frequency and quantity per attempt. Notice that when the quantity per attempt drops below 1 kg (Strategy 4.1), the quantity estimators and process state monitors drop out of the diversion path. This reflects the assumption that their detection threshold is somewhat less than 1 kg. The "Do Not Try" strategy is not shown in Table 7. However, it will be included in the model output in Section 3.

DETECTION PROBABILITIES

The influence diagram in Fig. 14 indicates that detection probabilities for each component depend on the adversary. Figure 15 shows the other events considered in the component detection assessment. The first node on the form is an adversary decision of whether or not to tamper with the component. For each component, we assume that adversaries will evaluate both options, and then make the decision that maximizes their expected utility.

TABLE 7. Diversion strategies.

Subsystem:	SS1--Electronic detection				SS2--Visual detection			SS3--Accounting				
Component:	1.1	1.2	1.3	1.4	2.1	2.2	2.3	3.1	3.2	Desired qty(kg)	Increment (kg)	Freq. (tries/y)	Description
Adversary strategy 1.1	0	0	0	0	2	2	0	0	0	10	10	1	Outsider armed attack
3.1	1	1	0	1	0	0	2	0	0	10	10	1	Normal hrs. diversion
3.2	1	1	1	1	2	2	0	0	0	10	10	1	Midnight diversion
3.3	1	0	1	0	2	2	0	0	0	10	10	1	Breach containment
3.4	0	0	0	0	2	2	0	0	0	10	10	1	Plant mgr. hostage
4.1	0	0	0	1	0	2	2	3	3	10	0.3	52	Normal hrs. small qty.
4.2	1	1	0	1	0	2	2	3	3	10	1	52	Normal hrs. large qty.
4.3	1	1	0	1	2	2	2	3	3	10	10	52	Normal hrs. whole thing
4.4	0	0	1	1	0	2	2	3	3	10	0.3	52	Midnight small qty.
4.5	1	1	1	1	0	2	2	3	3	10	1	52	Midnight large qty.
9.1	0	0	0	1	0	0	2	3	3	0.3	0.3	52	During wrk. hrs. from sampler
9.2	0	0	0	1	0	0	2	3	3	0.3	0.01	52	Small amounts from sampler
9.3	0	0	1	1	0	2	0	3	3	0.3	0.3	52	Midnight from sampler
11.1	0	0	0	0	2	2	0	3	3	10	10	1	Disable MC&A
11.2	1	1	0	1	0	2	2	3	3	10	10	1	Normal operation
11.3	1	1	1	1	2	2	0	3	3	10	10	1	Midnight
12.1	0	0	0	0	2	2	0	3	3	10	1	12	Disable MC&A
12.2	1	1	0	1	0	2	2	3	3	10	1	12	Normal operation
12.3	0	0	0	0	2	2	0	3	3	10	10	12	One shot disable MC&A
12.4	1	1	0	1	0	2	2	3	3	10	10	12	One shot normal oper.

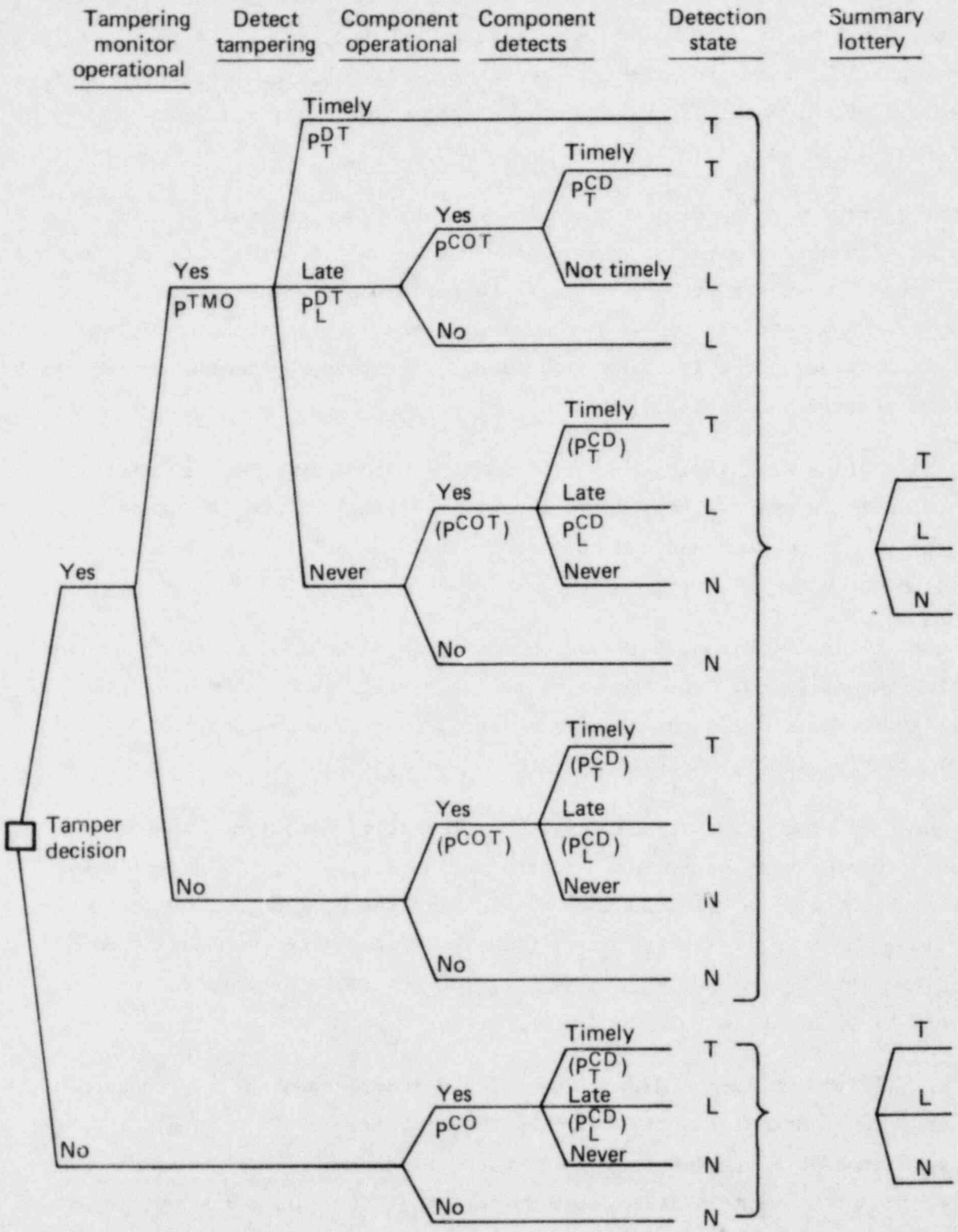


FIG. 15. Probability tree for detection assessment.

If the adversary tampers, and if there are tamper monitors on the component, then the first event considered is whether or not the tamper monitors are operational. If they are, then the adversary will be detected timely, late, or never. If they are not operational, then the adversary will not be detected.

Tampering takes on a special meaning in the case of Subsystem 2. There, the detectors are guards or other employees. We define tampering for Subsystem 2 as an attempt by the adversary to gain the collusion of noninvolved employees. An "operational" tamper system means that the noninvolved employee is aware that he or she is being approached. The tamper detection means the employee reports the contact.

The next node in Fig. 15 is whether or not the component is operational. Given that it is operational, detection will be timely, late, or never. If both the tamper monitor and the component itself detect late, the overall detection is late.

With the no-tamper option, only the operation and detection by the component itself are considered. The braces on the right in Fig. 15 show that the lotteries for both decisions can be collapsed to three-branch detection nodes: timely, late, and never.

We assume that the adversary chooses the alternative with the lowest probability of timely detection. If the probability of timely detection is zero, as with records systems, then we minimize the probability of late detection. When there are repeated attempts, with nonzero probabilities of timely and late detection, we assume that the adversary maximizes the probability of never being detected.

Figure 16 is the influence diagram used for detection probability assessment. Five important assumptions are shown in the diagram:

- Component operation is influenced by tampering.
- The probability of component operation is not influenced by tamper detection, but only by the act of tampering itself.
- Component detection depends only on the component being operational, not on adversary tampering.

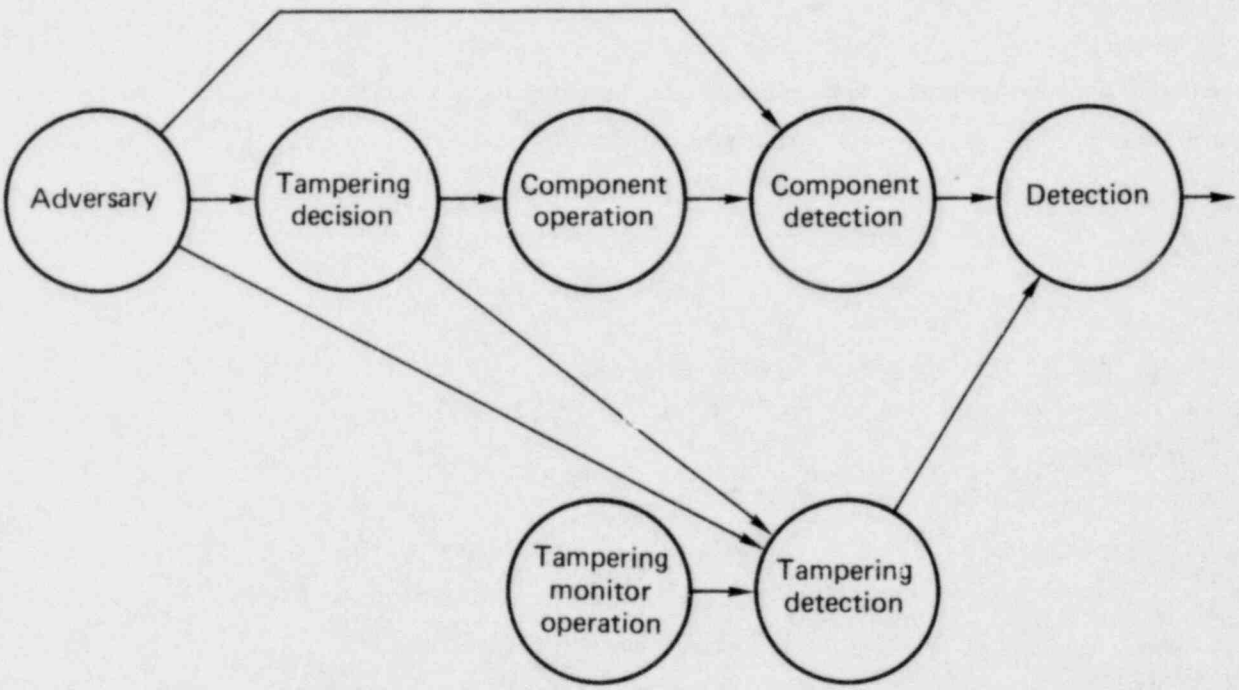


FIG. 16. Influence diagram for detection probability assessment.

- Detection probability varies with the adversary. We assume that SNM will be acquired by all adversaries in sufficient quantity to trigger detection.
- If the adversary attempts to steal a quantity that is below the detection thresholds, the component is not in the diversion path.
- The detection signal is treated the same whether it comes via the tamper monitor or the component itself.

The assessment of detection probabilities in Fig. 15 requires seven numbers: P^{TMO}, P_T^{DT}, P_L^{DT}, P^{COT}, P^{CO}, P_T^{CD}, and P_L^{CD} for each component and each adversary. These are listed in Table 8, along with comments on assessments.

DIVERSION TACTICS

The model in its general form allows the adversary to make a tactical decision on whether or not to abort an attempt at three points, that is, after each of the three subsystems detects the attempt. However, detection for the adversary means timely detection, and currently only two subsystems (material control and physical security) can detect in a timely fashion.

This means there are only six possible tactics:

Tactic	Description
0	Do not try
1	Always abort if detected
2	Abort only if SS1 detects
3	Abort only if SS2 detects
4	Abort only if SS1 and SS2 detect
5	Never abort.

See Appendix A for further discussion of tactics.

IDENTIFICATION

Figure 14 indicates that identification probabilities are dependent on the adversary, detection by subsystems, and on the adversary's abort tactics. Table 9 shows the illustrative assignments for the probabilities of timely and late identification.

TABLE 8a. Detection probability assessments.

Adversary/ component	Comments	Input Probabilities: Tamper monitor P^{TMO}	Tamper detection P_T^{DT}	Tamper detection P_L^{DT}	Component operation P^{COT}	Component operation P^{CO}	Component detection P_T^{CD}	Component detection P_L^{CD}	Computed probabilities: Tamper P(T)	Tamper P(L)	No tamper P(T)	No tamper P(L)	Tamper Decision
1/all components	Assume detection						1.0	0			1.0	0	No
3/1.1		0.9	0.7	0	0.3	0.9	0.7	0	0.71	0	0.63	0	No
3/1.2		0.9	0.7	0	0.3	0.9	0.7	0	0.71	0	0.63	0	No
3/1.3		0.9	0.7	0	0.3	0.9	0.9	0	0.78	0	0.81	0	Yes
3/1.4		0.9	0.9	0	0.1	0.9	0.95	0	0.83	0	0.86	0	Yes
3/2.1		0.5	0.2	0.7	0.2	0.8	0.5	0.3	0.2	0.35	0.4	0.24	Yes
3/2.2		0.8	0.5	0.1	0.5	0.9	0.5	0	0.55	0.02	0.45	0	
3/2.3		0.8	0.5	0.1	0.5	0.9	0.9	0	0.67	0.04	0.81	0	Yes
3/3.1		0.5	0	0.5	0.3	0.7	0	0.7	0	0.42	0	0.49	Yes
3/3.2		0.9	0	0.5	0.5	0.9	0	0.95	0	0.72	0	0.86	Yes
4/1.1		0.9	0.7	0	0.3	0.9	0.5	0	0.69	0	0.45	0	
4/1.2		0.9	0.7	0	0.3	0.9	0.5	0	0.69	0	0.45	0	
4/1.3		0.9	0.7	0	0.3	0.9	0.9	0	0.78	0	0.81	0	Yes
4/1.4		0.9	0.9	0	0.1	0.9	0.95	0	0.83	0	0.86	0	Yes
4/2.1		0.5	0.2	0.7	0.3	0.5	0.3	0.3	0.18	0.37	0.15	0.15	
4/2.2		0.8	0.5	0.1	0.5	0.9	0.2	0	0.46	0.03	0.18	0	
4/2.3		0.8	0.5	0.1	0.5	0.9	0.5	0	0.55	0.02	0.45	0	

(continued)

TABLE 8b. Detection probability assessments (continued).

Adversary/ component	Comments	Input Probabilities: Tamper monitor P^{TMO}	Tamper detection P_T^{DT}	Tamper detection P_L^{DT}	Component operation P^{COT}	Component operation P^{CO}	Component detection P_T^{CD}	Component detection P_L^{CD}	Computed probabilities: Tamper P(T)	Tamper P(L)	No tamper P(T)	No tamper P(L)	Tamper Decision
4/3.1		0.5	0	0.5	0.3	0.7	0	0.7	0	0.42	0	0.49	Yes
4/3.3		0.9	0	0.5	0.5	0.9	0	0.7	0	0.66	0	0.63	
9/1.1	Assume not					0.9	0.8	0			0.72	0	
9/1.1	skilled enough					0.9	0.8	0			0.72	0	
9/1.3	to tamper					0.9	0.9	0			0.81	0	
9/1.4	with SS1					0.9	0.95	0			0.86	0	
9/2.1		0.7	0.5	0.3	0.5	0.5	0.5	0.3	0.44	0.48	0.25	0.15	
9/2.2		0.8	0.5	0.3	0.5	0.5	0.2	0	0.44	0.42	0.1	0	
9/2.3	P(T) assigned directly										0.6	0	
9/3.1	Assigned. Compare to adv. 4										0	0.7	
9/3.2	Same adv. 4										0	0.63	
11/1.1	Assume no tampering.					0.9	0.7	0			0.63	0	
11/1.2	If adv 11 or 12 tamper, then the component is <u>not</u> in Monitor Target					0.9	0.7	0			0.63	0	
11/1.3	Set, beause adv. 11 & 12 have authority to					0.9	0.9	0			0.81	0	
11/1.4	disable Subsystem 1.					0.9	0.95	0			0.86	0	

(continued)

TABLE 8c. Detection probability assessments (concluded).

Adversary/ component	Comments	Input probabilities: Tamper monitor P^{TMO}	Tamper detection P_T^{DT}	Tamper detection P_L^{DT}	Component operation P^{COT}	Component operation P^{CO}	Component detection P_T^{CD}	Component detection P_L^{CD}	Computed probabilities: Tamper P(T)	Tamper P(L)	No tamper P(T)	No tamper P(L)	Tamper Decision
11/2.1		0.5	0.2	0.7	0.2	0.5	0.5	0.3	0.14	0.82	0.25	0.15	Yes
11/2.2	Assume no tamper					0.9	0.5	0			0.45	0	Yes
11/2.3	Assume collusion	0.8	0.2	0.2	0.4	0.7	0.5	0.3	0.29	0.33	0.35	0.21	Yes
11/3.1	Assigned similar to adv. 3							0	0	0.4	0	0.5	Yes
11/3.2	Assigned similar to adv. 3							0	0	0.7	0	0.8	
12/1.1	Same as adv. 11 max. P(never) where indicated:										0.63	0	
12/1.2											0.63	0	
12/1.3											0.81	0	
12/1.4											0.86	0	
12/2.1	max. P(never)								0.14	0.82	0.25	0.15	
12/2.2											0.45	0	
12/2.3									0.29	0.33	0.35	0.21	
12/3.1									0	0.4	0	0.5	Yes
12/3.2									0	0.7	0	0.8	Yes

TABLE 9. Identification probability assessment.

Detection by subsystems			Abort decision	Adversary identification probability											
				1		3		4		9		11		12	
1	2	3		T	L	T	L	T	L	T	L	T	L	T	L
T	-	-	Y	1.0	0	0.3	0.3	0.2	0.2	0.5	0.3	0.1	0.1	0.1	0.1
T	T	-	Y	1.0	0	0.5	0.4	0.5	0.4	0.6	0.2	0.2	0.2	0.2	0.2
T	T	L	N	1.0	0	0.8	0.1	0.8	0.1	0.8	0.1	0.5	0.1	0.5	0.1
T	T	N	N	1.0	0	0.8	0.05	0.8	0.05	0.8	0.05	0.5	0.1	0.5	0.1
T	L	L	N	1.0	0	0.6	0.3	0.5	0.3	0.5	0.3	0.3	0.2	0.3	0.2
T	N	L	N	1.0	0	0.6	0.2	0.5	0.2	0.5	0.2	0.3	0.1	0.3	0.1
T	L	N	N	1.0	0	0.6	0.3	0.5	0.3	0.5	0.3	0.3	0.2	0.3	0.2
T	N	N	N	1.0	0	0.6	0.1	0.5	0.1	0.5	0.1	0.3	0.1	0.3	0.1
N	T	-	Y	0	0	0.5	0.3	0.5	0.3	0.5	0.3	0.3	0.2	0.3	0.2
N	T	L	N	0	0	0.8	0.1	0.7	0.1	0.7	0.1	0.5	0.1	0.5	0.1
N	T	N	N	0	0	0.8	0.05	0.7	0.05	0.7	0.05	0.5	0.05	0.5	0.05
N	L	L	N	0	0	0	0.5	0	0.3	0	0.3	0	0.2	0	0.2
N	L	N	N	0	0	0	0.3	0	0.2	0	0.2	0	0.1	0	0.1
N	N	N	N	0	0	0	0	0	0	0	0	0	0	0	0

T = Timely

L = Late

N = Never or no

Y = Yes

- = Not applicable

INTERRUPTION

The probability that the guards will interrupt the sequence depends on the adversary and timely identification. The illustrative probabilities are shown in Table 10.

ADVERSARY UTILITIES

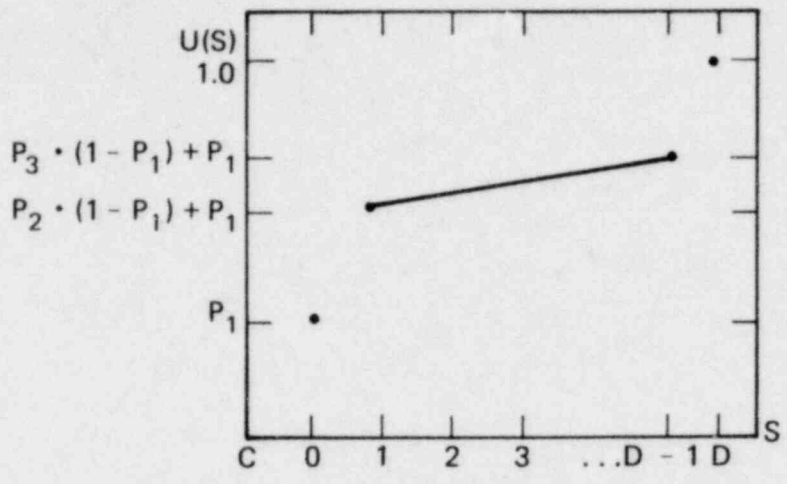
Figure 17 is a copy of the form used for assigning adversary utility functions. Three input numbers are required: P_1, P_2, P_3. These are tabulated for all adversaries in Table 11. The forms of the curves are shown in Fig. 18.

TABLE 10. Interruption probabilities.

Adversary	P(Interruption | Timely identification)
1	0.20
3	0.99
4	0.99
9	0.99
11	0.99
12	0.99

TABLE 11. Adversary utility function parameters.

Adversary case	Name	P_1	P_2	P_3	a	b
1.0	Base case--adv. 1	0.6	0	0	0	0.6
3.0	Base case--adv. 3	0.7	0	0	0	0.7
3.1	Low capture aversion	0.2	0	0	0	0.2
3.2	High capture aversion	0.9	0	0	0	0.9
4.0	Base case--adv. 4	0.7	0.2	0.5	0.09/(D-2)	0.76-a
4.1	Low capture aversion	0.2	0.2	0.5	0.24/(D-2)	0.36-a
4.2	High capture aversion	0.9	0.2	0.5	0.03/(D-2)	0.92-a
4.3	Valuable first increment	0.5	0.2	0.25	0.03/(D-2)	0.60-a
4.4	Valuable last increment	0.5	0.01	0.05	0.02/(D-2)	0.51-a
9.0	Base case--adv. 9	0.7	0.3	0.9	0.18/(D-2)	0.79-a
11.0	Base case--adv. 11	0.9	0	0	0	0.9
12.0	Base case--adv. 12	0.9	0.6	0.9	0.03/(D-2)	0.96-a



C = Capture state
S = Number of successes
D = Number of desired successes

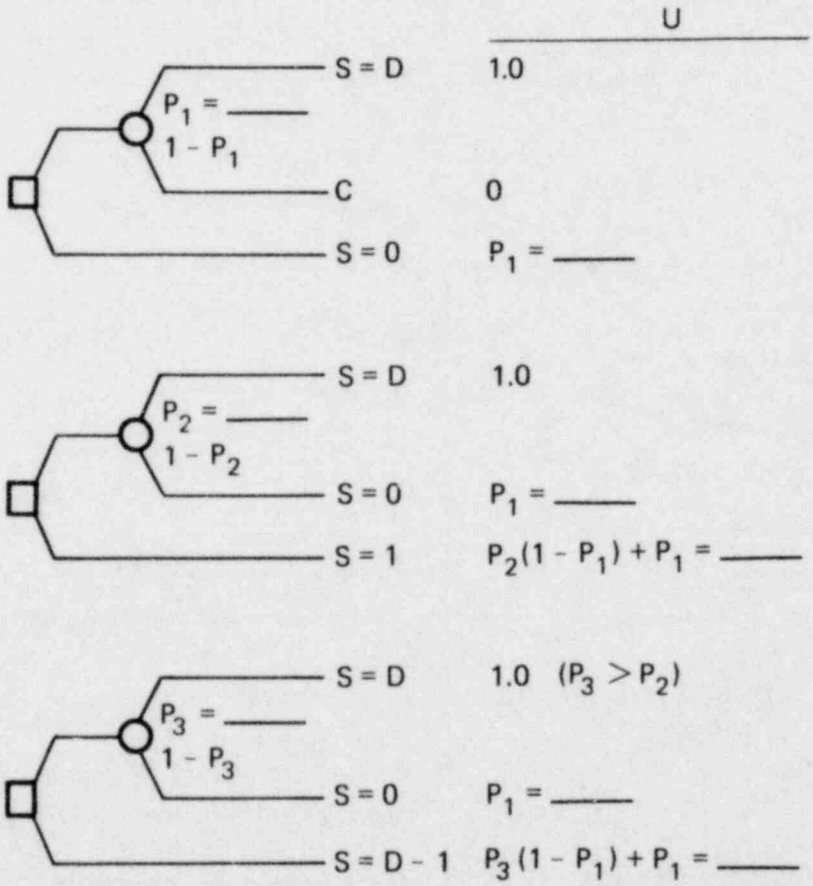


FIG. 17. Adversary utility function assessment form.

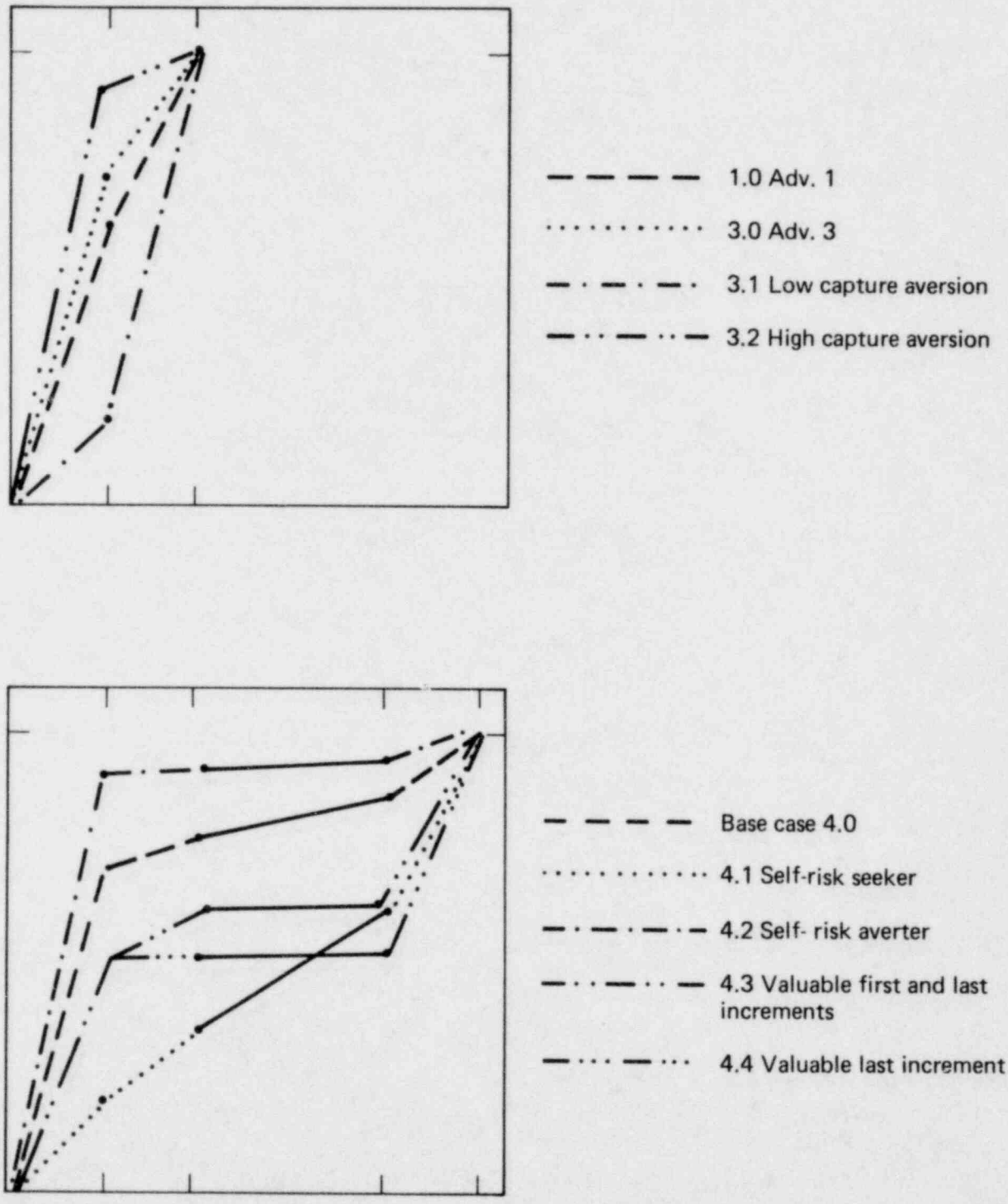


FIG. 18. Forms of selected adversary utility functions.

III. ILLUSTRATIVE ANALYSIS

INTRODUCTION

This section demonstrates types of analysis facilitated by the model using illustrative data. Because of the nature of the data, we do not draw substantive conclusions. However, we point out general insights to improve the user's understanding of the model and its sensitivity to some key parameters. Together with data from an actual facility, these insights would give a decision maker a better understanding of the safeguards system's effectiveness against various types of adversaries.

We shall present four general types of model output:

- Adversary decision analysis
- Safeguards evaluation
- System performance graphs
- Sensitivity studies.

The first output evaluates the strategies and tactics available to each adversary. For a given system design, the model examines every alternative the adversary may consider and calculates the adversary's utility and other parameters such as detection probabilities.

The safeguards design evaluation produces an aggregated form of output. It repeats the adversary decision process for each system design, allows each adversary to choose his or her best alternative, and then produces system performance measures aggregated across all adversary types. This type of analysis focuses on the benefits of one design, or one performance standard, over another.

The final type of analysis, sensitivity analysis, demonstrates how the model results change when input data are changed. The example we give shows the results of changing adversary utility functions and systems designs. Numerous examples of other sensitivities can also be examined, and are given in Ref. 6.

ADVERSARY DECISION ANALYSIS

Adversary Tactical Decisions

Table 12a shows the most detailed level of model output for:

- System Design 1--testbed design (See Table 5)
- Adversary 1--outsiders
- Diversion Strategy 1--armed attack (See Table 7).

Table 12a tabulates eight output parameters for all six tactics Adversary 1 could choose. These parameters are:

- E(U) Adversary expected utility
- P_{SD} Probability of successfully diverting the desired quantity of SNM (10 kg for Adversary 1)
- E(N) Expected number of attempts by the adversary
- E(SNM) Expected quantity of SNM diverted
- E(S) The expected number of successes
- P_D Probability of detection on each attempt
- P_C Probability that the series of attempts will end in capture
- P_{ID} Probability that the adversary will be identified during the attempt(s).

Five tactics, ranging from "always abort if detected by any subsystem" to "never abort," are shown in Table 13. In addition, tactic 0 represents the situation in which the adversary does not initiate the diversion.

These output measures indicate what is happening during each tactic and give insight into how a tactical decision is made. Look at the results in Table 12a. The physical security force (Subsystem 2) will always detect Adversary 1 (see Table 8). Tactics 1 and 3 are to abort if detected by the guards. Because guard detection is certain, Tactics (T) 1 and 3 will always mean failure: no SNM. Therefore, these have low adversary utility. With T2, T4, and T5, the adversary has 0.56 probability of success and 0.2 probability of capture. Detection by Subsystem 1 (SS1), the monitors, is superfluous, so T2, T4, and T5 are the same.

TABLE 12a. Example evaluation of tactics--outsider.

SYSTEM DESIGN 1- TESTBED; ADVERSARY 1; DIVERSION STRATEGY: ASSAULT

TACTIC	E(U)	P(SD)	E(N)	E(S)	E(SNM)	P(D)	P(C)	P(ID)
0	.600	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1	.480	0.000	1.000	0.000	0.000	1.000	.200	1.000
2	.704	.560	1.000	.560	5.600	1.000	.200	1.000
3	.480	0.000	1.000	0.000	0.000	1.000	.200	1.000
4	.704	.560	1.000	.560	5.500	1.000	.200	1.000
5	.704	.560	1.000	.560	5.600	1.000	.200	1.000

TABLE 12b. Example evaluation of tactics--insider.

SYSTEM DESIGN 1- TESTBED; ADVERSARY 4; DIVERSION STRATEGY: NORMAL,.3K

TACTIC	E(U)	P(SD)	E(N)	E(S)	E(SNM)	P(D)	P(C)	P(ID)
0	.700	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1	.007	.000	3.943	.221	.066	.906	.990	1.000
2	.007	.000	3.804	.383	.115	.906	.990	1.000
3	.007	.000	2.138	.505	.152	.906	.990	1.000
4	.007	.000	2.079	.546	.164	.906	.990	1.000
5	.008	.000	1.714	.580	.174	.906	.990	1.000

TABLE 13. Adversary tactics.

Tactic number	Abort if detected by subsystem:
0	Don't try at all
1	1 or 2
2	1 only
3	2 only
4	1 and 2
5	Never

The two lotteries represented by T1 and T5 are shown in Fig. 19. The calculation of expected utility is also shown. The outcome states are those shown in Fig. 10; adversary utilities are from Table 11. The expected utilities for each tactic are shown in ovals, and they agree with those tabulated in Table 12a. The best choice is to "never abort." To make the attempt but later abort is worse than not trying at all, because of the 0.2 probability of capture. The 0.56 probability of success comes from the 0.7 probability of acquiring SNM (Table 6) times (1.0 - 0.2 = 0.8) probability of no capture (Table 10).

Table 12a is a very simple case. Table 12b, on the other hand, describes a more complex case. The situation is:

- Design 1--testbed
- Adversary 4--insider, multiple attempts possible
- Diversion Strategy 3--steal SNM in one 10-kg quantity.

Diverting one large quantity is somewhat simpler than an adversary stealing many small increments.

If the adversary makes a try, his or her utility shows T5--"never abort"--to be the best tactic because it has the highest value for expected SNM diverted and also the lowest probability of capture. T1 and T2 have significantly lower expected utility. This decreased utility depends on several factors. The expected amount of SNM diverted is smaller because so many of the tries are aborted. Aborting the tries decreases the probability of capture; however, because the number of tries goes up, the probability of eventually being captured goes up too. Therefore, T5 is the best tactic. Nevertheless, the utility of T5 is less than the utility of not trying at all. The adversary will have the highest expected utility if he or she does nothing.

Adversary Strategic Decisions--General Description

Table 14 evaluates all tactics for every diversion strategy (DS) considered by Adversary 3. Table 7 showed these strategies:

- DS1 Normal hours diversion
- DS2 Midnight diversion

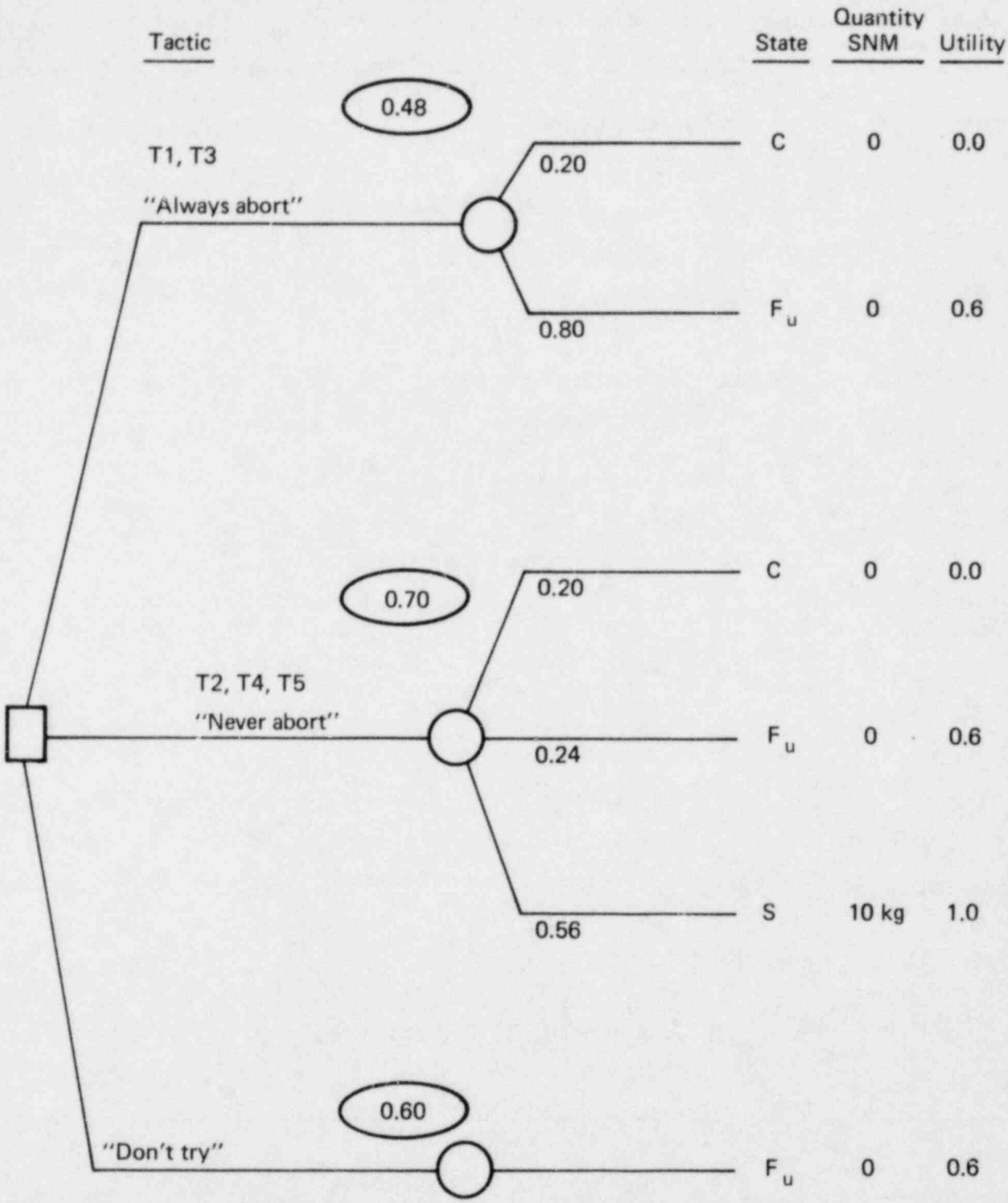


FIG. 19. Adversary 1's tactical decision problem.

TABLE 14. Adversary 3's strategic decision evaluation.

SYSTEM DESIGN 1- TESTBED; ADVERSARY 3; DIVERSION STRATEGY: NORMAL,10K

TACTIC	E(U)	P(SD)	E(N)	E(S)	E(SNM)	P(D)	P(C)	P(ID
0	.700	0.000	0.000	0.000	0.000	0.000	0.000	0.00
1	.493	.004	1.000	.004	.043	.992	.298	.30
2	.492	.013	1.000	.013	.130	.992	.302	.30
3	.372	.126	1.000	.126	1.259	.992	.523	.52
4	.370	.131	1.000	.131	1.306	.992	.528	.53
5	.261	.222	1.000	.222	2.223	.992	.722	.72

SYSTEM DESIGN 1- TESTBED; ADVERSARY 3; DIVERSION STRATEGY: MIDNITE,10

TACTIC	E(U)	P(SD)	E(N)	E(S)	E(SNM)	P(D)	P(C)	P(ID
0	.700	0.000	0.000	0.000	0.000	0.000	0.000	0.00
1	.494	.002	1.000	.002	.023	.996	.296	.29
2	.493	.003	1.000	.003	.029	.996	.296	.29
3	.382	.274	1.000	.274	2.741	.996	.572	.57
4	.382	.274	1.000	.274	2.742	.996	.572	.57
5	.347	.295	1.000	.295	2.951	.996	.631	.63

SYSTEM DESIGN 1- TESTBED; ADVERSARY 3; DIVERSION STRATEGY: BREACH,10K

TACTIC	E(U)	P(SD)	E(N)	E(S)	E(SNM)	P(D)	P(C)	P(ID
0	.700	0.000	0.000	0.000	0.000	0.000	0.000	0.00
1	.515	.037	1.000	.037	.375	.935	.281	.28
2	.514	.047	1.000	.047	.465	.935	.286	.28
3	.414	.297	1.000	.297	2.973	.935	.536	.54
4	.412	.300	1.000	.300	3.002	.935	.540	.54
5	.381	.324	1.000	.324	3.241	.935	.595	.60

SYSTEM DESIGN 1- TESTBED; ADVERSARY 3; DIVERSION STRATEGY: HOSTAGE,10

TACTIC	E(U)	P(SD)	E(N)	E(S)	E(SNM)	P(D)	P(C)	P(ID
0	.700	0.000	0.000	0.000	0.000	0.000	0.000	0.00
1	.804	.577	1.000	.577	5.766	.200	.099	.10
2	.791	.673	1.000	.673	6.733	.200	.158	.16
3	.804	.577	1.000	.577	5.766	.200	.099	.10
4	.791	.673	1.000	.673	6.733	.200	.158	.16
5	.791	.673	1.000	.673	6.733	.200	.158	.16

- DS 3 Breach containment
- DS 4 Take plant manager hostage.

The adversary only tries once. The system design is System 1--testbed design.

With Diversion Strategy DS1, Tactic 5 has the highest expected value for SNM diverted--E(SNM)--and a high probability of capture--P_C. The high P_C relative to T4 outweighs the increase in E(SNM) relative to T4, so T5 has lower expected utility than T4.

(The probability of detection is the same for all tactics under any given strategy. This is because tactics only influence responses after detection. The probabilities of final identification and capture vary because some tactics cause more aborted attempts than others.)

Adversary 3 is nearly certain to be detected, but has rather low probabilities of being identified (See Table 9), especially if he or she aborts (0.3 if abort given SS1 detection, 0.6 if no abort). Notice that the expected utility for DS1, given an attempt, is E(U) = 0.493, which is less than the expected utility of not trying: 0.7.

With Diversion Strategy DS2--midnight diversion--the two person rule is not in the diversion path, but the personnel monitors are (Table 7). Notice with T3, T4, and T5, the E(SNM) is 100 times greater than for T1 and T2. However, P_C is almost double; thus, the adversary's utility is lower for T5 than for T1 and T2. Because T1 and T2 have the same probability of capture, they yield about the same expected utility. The choice between T1 and T5 is shown in Fig. 20. Once again, "not attempting" is the best strategy.

With Diversion Strategy DS3--breach containment--we assume that the personnel monitors and process state monitors drop out of the Monitor Target Sets. In this case, the detection probability drops, E(SNM) is slightly higher, and P_C is slightly lower, so E(U) increases.

Diversion Strategy DS4--taking the plant manager hostage--shows that the adversary should make the attempt. All of the tactics have expected utilities higher than the utility of not trying. Tactics 1 and 3, however, have the

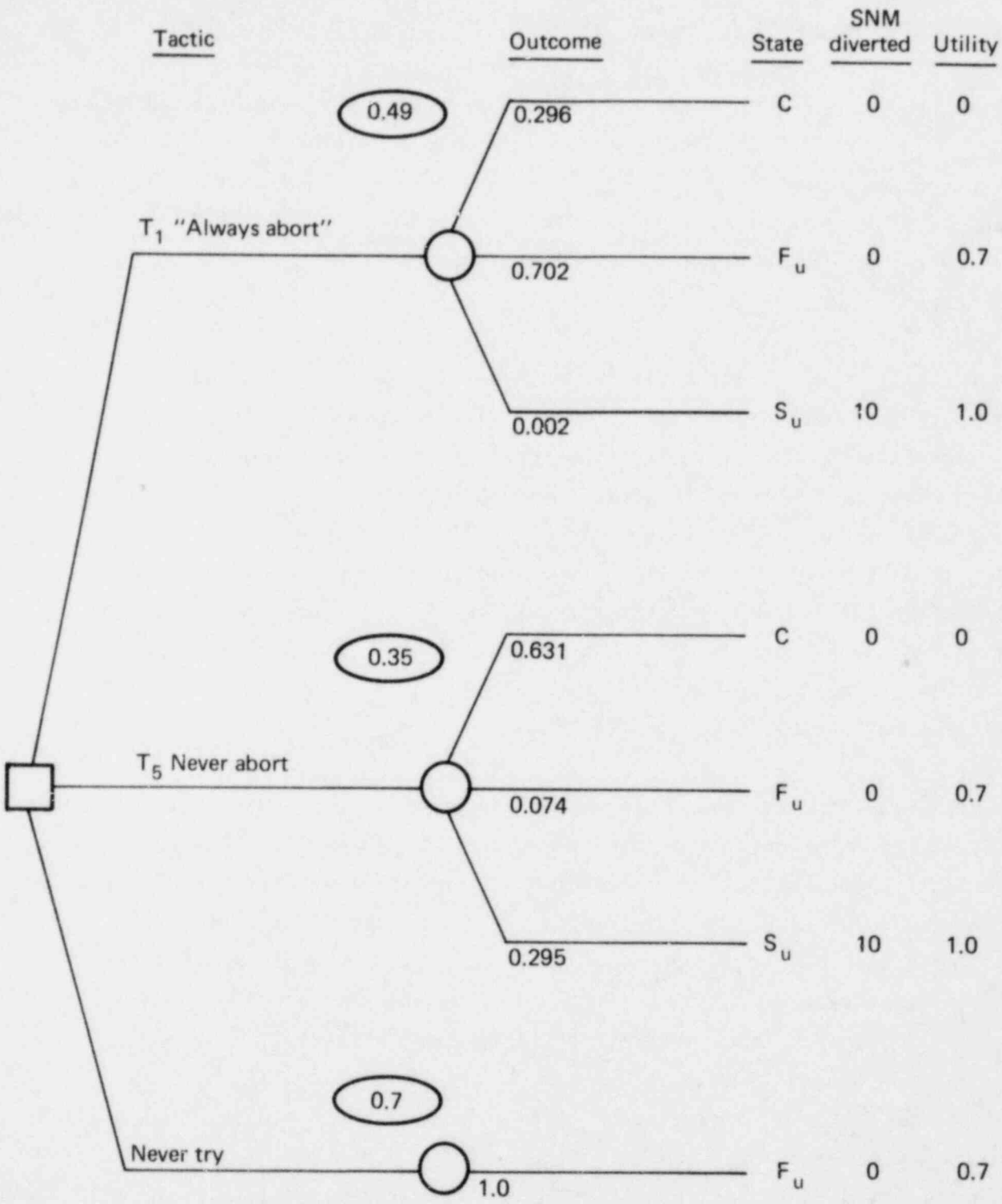


FIG. 20. Adversary 3's choice of tactics in diversion strategy 2.

highest utilities. Both of these involve aborting if the physical security force detects the try. Even though less SNM is expected to be diverted with T1 and T3, that probability of capture decreases when the adversary chooses these two tactics. Thus, the adversary should make the try but abort if he or she is detected by the guards. This last strategy has low P_{ID} even when the adversary does not abort. This is a result of the input data, which assume that Adversary 3 is sophisticated and hard to identify, regardless of diversion strategy.

Table 15 shows the best tactic for each diversion strategy. This makes it easy to examine each adversary's choice of diversion strategy number. As we have discussed so far, for System Design 1 Adversary 1's best tactic is T5. The best tactic for each of Adversary 3's four strategies are given next. Recall the "0" strategy corresponds to "no attempt."

TABLE 15. Evaluation of all adversaries' strategies--design 1.

SYSTEM DESIGN 1; C(S) 6.6; C(D) 5.8; C(T) 12.4

ADV.STRAT		TACTIC	E(U)	P(SD)	E(N)	E(S)	E(SNM)	P(D)	P(C)	P(ID)
1.0		0	.600	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1.1	*	5	.704	.560	1.000	.560	5.600	1.000	.200	1.000
3.0		0	.600	0.000	0.000	0.000	0.000	0.000	0.000	0.000
3.1		1	.493	.004	1.000	.004	.043	.992	.298	.301
3.2		1	.494	.002	1.000	.002	.023	.996	.296	.299
3.3		1	.515	.037	1.000	.037	.375	.935	.281	.284
3.4	*	3	.804	.577	1.000	.577	5.766	.200	.099	.100
4.0	*	0	.700	0.000	0.000	0.000	0.000	0.000	0.000	0.000
4.1		5	.008	.000	1.714	.580	.174	.906	.990	1.000
4.2		5	.008	.000	1.612	.497	.497	.972	.990	1.000
4.3		5	.312	.311	1.379	.311	3.113	.976	.687	.694
4.4		5	.008	.000	1.600	.488	.146	.979	.990	1.000
4.5		5	.008	.000	1.580	.472	.472	.994	.990	1.000
9.0	*	0	.700	0.000	0.000	0.000	0.000	0.000	0.000	0.000
9.1		5	.059	.053	1.522	.053	.016	.944	.938	.948
9.2		5	.007	.000	1.550	.056	.001	.944	.990	1.000
9.3		5	.100	.094	1.933	.094	.028	.973	.898	.907
11.0		0	.700	0.000	0.000	0.000	0.000	0.000	0.000	0.000
11.1	*	3	.929	.659	1.000	.659	6.594	.140	.042	.042
11.2		1	.812	.010	1.000	.010	.098	.986	.099	.100
11.3		1	.811	.001	1.000	.001	.011	.999	.099	.100
12.0	*	0	.900	0.000	0.000	0.000	0.000	0.000	0.000	0.000
12.1		5	.094	.086	5.296	3.445	3.445	.250	.905	.915
12.2		5	.010	.001	2.303	1.050	1.050	.988	.990	.999
12.3		5	.837	.837	2.036	.837	8.371	.250	.163	.164
12.4		5	.569	.569	1.701	.569	5.685	.988	.431	.435

Adversary 4's Strategies

Adversary 4 can make repeated tries on taking increments of differing sizes to accumulate 10 kg. The strategies are:

- DS0 Do not attempt to divert SNM
- DS1 Normal hours, 0.3 kg increment
- DS2 Normal hours, 1.0 kg increment
- DS3 Normal hours, 10.0 kg increment
- DS4 Midnight diversion, 0.3 kg increment
- DS5 Midnight diversion, 1.0 kg increment.

The best strategy, as shown in Table 15, is DS0. For DS1 through DS5, the probabilities of detection, identification, and capture are all relatively high, so Adversary 4 has a difficult tradeoff between capture and getting the desired SNM. Though Table 15 does not show this, all tactics have about the same expected utility for strategies DS1, DS2, DS4, and DS5. However, DS3--taking the one large quantity--is preferred to the strategies of taking small increments. Notice that for DS1 through DS5, the adversary should minimize the number of tries; minimizing P_D is not optimal. However, none of the strategies has a high enough utility to induce the adversary to try.

If a system can be designed so that an adversary is better off not making an attempt, the adversary effectively has been deterred. Adversary expected utility therefore serves as a measure of deterrence. If no attempt is made, one might postulate that the system is safe enough to protect against Adversary 4.

Table 16 shows the same type of results for System Design 2 that Fig. 14 showed for System 1. Results for Adversary 4 show that the two diversion strategies that involve 0.3-kg increments have the same expected utility as do the diversions involving 1.0-kg increments. The adversary's choice can be pictured as a decision tree, shown in Fig. 21. For each strategy, the outcome probabilities and adversary preferences are assigned. The expected utility for each strategy is calculated, showing that the "one large quantity" strategy is best. We could construct a similar decision tree for each adversary's choice of diversion path as well as quantity per try.

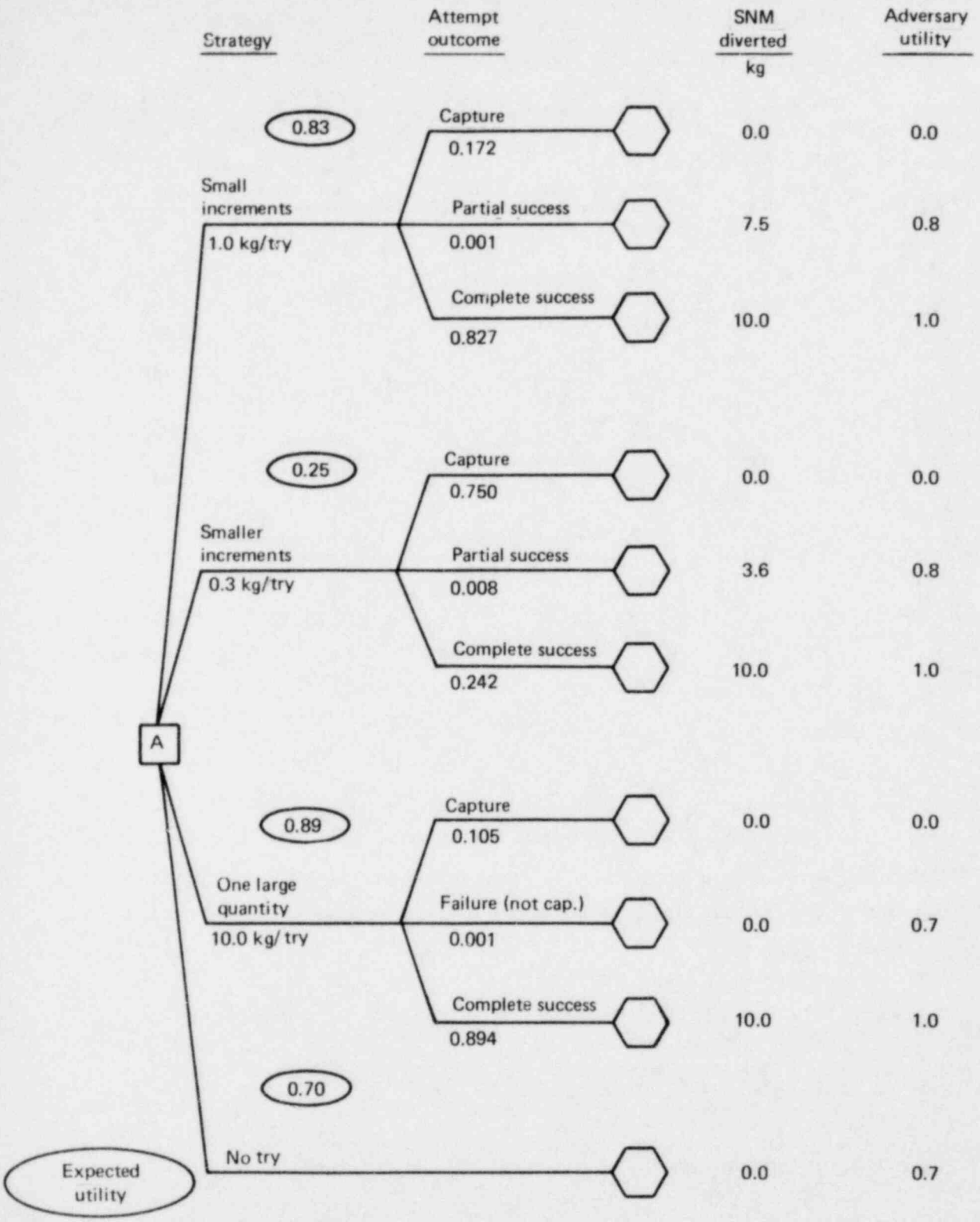


FIG. 21. Sample strategic choice by adversary 4 in design 2.

TABLE 16. Evaluation of all adversaries' strategies--design 2.

SYSTEM DESIGN 2- MINIMAL; C(S) .5; C(D) 21.5; C(T) 22.0

ADV.	STRATEGY	TACTIC	E(U)	P(SD)	E(N)	E(S)	E(SNM)	P(D)	P(C)	P(ID)
1	DETERRED	0	.600	0.000	0.000	0.000	0.000	0.000	0.000	0.00
1	ASSAULT *	5	.704	.560	1.000	.560	5.600	1.000	.200	1.00
3	DETERRED	0	.700	0.000	0.000	0.000	0.000	0.000	0.000	0.00
3	NORMAL,10K *	5	.940	.800	1.000	.800	8.000	0.000	0.000	0.00
3	MIDNITE,10	3	.804	.577	1.000	.577	5.766	.200	.099	.10
3	BREACH,10K	3	.804	.577	1.000	.577	5.766	.200	.099	.10
3	HOSTAGE,10	3	.804	.577	1.000	.577	5.766	.200	.099	.10
4	DETERRED	0	.700	0.000	0.000	0.000	0.000	0.000	0.000	0.00
4	NORMAL,.3K	5	.248	.242	26.279	20.231	6.069	0.000	.750	.75
4	NORMAL,1K	5	.828	.827	12.944	9.563	9.563	0.000	.172	.17
4	NORMAL,10K *	3	.894	.894	2.305	.894	8.942	.150	.105	.10
4	MIDNITE,.3	5	.248	.242	26.279	20.231	6.069	0.000	.750	.75
4	MIDNITE,1K	5	.828	.827	12.944	9.563	9.563	0.000	.172	.17
9	DETERRED	0	.700	0.000	0.000	0.000	0.000	0.000	0.000	0.00
9	SAMPLER,.3 *	5	.816	.815	9.143	.815	.245	0.000	.183	.18
9	SAMPLER.01	5	.008	.000	22.286	2.130	.021	0.000	.990	1.00
9	MIDNITE,.3	5	.816	.815	9.143	.815	.245	0.000	.183	.18
11	DETERRED	0	.900	0.000	0.000	0.000	0.000	0.000	0.000	0.00
11	KILLMCA,10	3	.929	.659	1.000	.659	6.594	.140	.042	.04
11	NORMAL,10K *	5	.980	.800	1.000	.800	8.000	0.000	0.000	0.00
11	MIDNITE,10	3	.929	.659	1.000	.659	6.594	.140	.042	.04
12	DETERRED	0	.900	0.000	0.000	0.000	0.000	0.000	0.000	0.00
12	KILLMCA,1K	5	.101	.092	5.559	3.655	3.655	.250	.899	.90
12	NORMAL,1K	5	.656	.653	11.320	8.264	8.264	0.000	.343	.34
12	KILLMCA,10	3	.864	.863	2.429	.863	8.632	.250	.136	.13
12	NORMAL,10K *	5	.998	.998	2.238	.998	9.980	0.000	.002	.00

Adversaries' 9, 11, and 12 Strategies

Table 15 shows that Adversary 9, with minor resources, has little chance of stealing even 0.3 kg. The adversary should maximize utility by not trying. In Table 16 under System Design 2, Adversary 9, like all adversaries, does make the attempt.

Adversaries 11 and 12 are plant managers. The diversion strategies for Adversary 11 are:

- DS1 Disable MC&A system (Subsystem 1)
- DS2 Divert as part of normal operations
- DS3 Midnight diversion.

Exploiting his or her authority over the MC&A system is the best strategy in Table 15. However, the best tactic in this case is to abort if the guards catch on to what is happening. With DS1, the perpetrator faces almost no risk (P_C = 0.042) and the SNM is "there for the taking."

Adversary 12 has the option of repeated attempts, either for one large quantity or small increments:

- DS1 Disable MC&A--small increments
- DS2 During normal operations--small increments
- DS3 Disable MC&A--large quantity
- DS4 Normal operations--one large quantity.

None of Adversary 12's strategies in Table 15 yield a utility high enough to warrant an attempt, although disabling the MC&A System comes close to being worth an attempt.

Since Adversaries 11 and 12 are both plant managers, with 12 having more flexibility than 11, Adversary 12's best option should have higher utility than Adversary 11's. However, comparing results in Table 15 shows that this is not the case. This is because the model forces Adversary 12 to make repeated tries if the first is unsuccessful. The first try has high utility: E(U) = 0.929. Unfortunately, after the first try there is an increasing chance of late detection, identification, and capture. This chain of events, combined with the adversary's utility, which values SNM very little relative to not being captured, causes the utility to drop.

The expected number of tries may be quite large, especially when the probability of detection is low. Such cases are evident in Table 16.

SAFEGUARDS DESIGN EVALUATION

Overview

Table 17 summarizes optimal adversary strategies and tactics for each system design. We will discuss the merits of individual designs in the next paragraphs. However, adversary-to-adversary comparisons are noteworthy:

- Regardless of the system, Adversary 1 will not be deterred; therefore, each system anticipates some risk of diversion.
- Given the frequency distribution over adversaries in Table 1, the greatest contribution to overall E(SNM) is by Adversary 4. Thus, any system that reduces Adversary 4's E(SNM) will have generally lower C_D (diversion cost).
- Adversary 11's E(SNM) changes drastically over designs 2, 5, and 7. This is because physical security varies, and the adversary is choosing different diversion strategies in response to the variations. System

TABLE 17. Adversary decisions for all system designs.

SYSTEM DESIGN 1- TESTBED; C(S) 6.6; C(D) 5.8; C(T) 12.5

ADV.STRATEGY	TACTIC	E(U)	P(SD)	E(N)	E(S)	E(SNM)	P(D)	P(C)	P(ID)
1 ASSAULT	5	.704	.560	1.000	.560	5.600	1.000	.200	1.000
3 HOSTAGE,10	3	.804	.577	1.000	.577	5.766	.200	.099	.100
4 DETERRED	0	.700	0.000	0.000	0.000	0.000	0.000	0.000	0.000
9 DETERRED	0	.700	0.000	0.000	0.000	0.000	0.000	0.000	0.000
11 KILLMCA,10	3	.929	.659	1.000	.659	6.594	.140	.042	.042
12 DETERRED	0	.900	0.000	0.000	0.000	0.000	0.000	0.000	0.000

SYSTEM DESIGN 2- MINIMAL; C(S) .5; C(D) 21.5; C(T) 22.0

ADV.STRATEGY	TACTIC	E(U)	P(SD)	E(N)	E(S)	E(SNM)	P(D)	P(C)	P(ID)
1 ASSAULT	5	.704	.560	1.000	.560	5.600	1.000	.200	1.000
3 NORMAL,10K	5	.940	.800	1.000	.800	8.000	0.000	0.000	0.000
4 NORMAL,10K	3	.894	.894	2.305	.894	8.942	.150	.105	.107
9 SAMPLER,.3	5	.816	.815	9.143	.815	.245	0.000	.183	.185
11 NORMAL,10K	5	.980	.800	1.000	.800	8.000	0.000	0.000	0.000
12 NORMAL,10K	5	.998	.998	2.238	.998	9.980	0.000	.002	.002

SYSTEM DESIGN 3- MAXIMAL; C(S) 12.0; C(D) 2.7; C(T) 14.7

ADV.STRATEGY	TACTIC	E(U)	P(SD)	E(N)	E(S)	E(SNM)	P(D)	P(C)	P(ID)
1 ASSAULT	5	.704	.560	1.000	.560	5.600	1.000	.200	1.000
3 DETERRED	0	.700	0.000	0.000	0.000	0.000	0.000	0.000	0.000
4 DETERRED	0	.700	0.000	0.000	0.000	0.000	0.000	0.000	0.000
9 DETERRED	0	.700	0.000	0.000	0.000	0.000	0.000	0.000	0.000
11 DETERRED	0	.900	0.000	0.000	0.000	0.000	0.000	0.000	0.000
12 DETERRED	0	.900	0.000	0.000	0.000	0.000	0.000	0.000	0.000

SYSTEM DESIGN 4- MEDIUM ; C(S) 1.7; C(D) 15.4; C(T) 17.1

ADV.STRATEGY	TACTIC	E(U)	P(SD)	E(N)	E(S)	E(SNM)	P(D)	P(C)	P(ID)
1 ASSAULT	5	.704	.560	1.000	.560	5.600	1.000	.200	1.000
3 HOSTAGE,10	3	.804	.577	1.000	.577	5.766	.200	.099	.100
4 NORMAL,10K	2	.719	.718	2.622	.718	7.181	.532	.280	.283
9 SAMPLER,.3	5	.816	.815	9.143	.815	.245	0.000	.183	.185
11 KILLMCA,10	3	.929	.659	1.000	.659	6.594	.140	.042	.042
12 DETERRED	0	.900	0.000	0.000	0.000	0.000	0.000	0.000	0.000

SYSTEM DESIGN 5- 2-MAN ; C(S) 2.2; C(D) 5.8; C(T) 8.0

ADV.STRATEGY	TACTIC	E(U)	P(SD)	E(N)	E(S)	E(SNM)	P(D)	P(C)	P(ID)
1 ASSAULT	5	.704	.560	1.000	.560	5.600	1.000	.200	1.000
3 HOSTAGE,10	3	.804	.577	1.000	.577	5.766	.200	.099	.100
4 DETERRED	0	.700	0.000	0.000	0.000	0.000	0.000	0.000	0.000
9 DETERRED	0	.700	0.000	0.000	0.000	0.000	0.000	0.000	0.000
11 KILLMCA,10	3	.929	.659	1.000	.659	6.594	.140	.042	.042
12 DETERRED	0	.900	0.000	0.000	0.000	0.000	0.000	0.000	0.000

(continued)

TABLE 17. (Continued)

SYSTEM DESIGN 6- R GUARD; C(S) 2.1; C(D) 2.7; C(T) 4.8

ADV.STRATEGY	TACTIC	E(U)	P(SD)	E(N)	E(S)	E(SNM)	P(D)	P(C)	P(ID)
1 ASSAULT	5	.704	.560	1.000	.560	5.600	1.000	.200	1.000
3 DETERRED	0	.700	0.000	0.000	0.000	0.000	0.000	0.000	0.000
4 DETERRED	0	.700	0.000	0.000	0.000	0.000	0.000	0.000	0.000
9 SAMPLER,.3	5	.816	.815	9.143	.815	.245	0.000	.183	.185
11 DETERRED	0	.900	0.000	0.000	0.000	0.000	0.000	0.000	0.000
12 DETERRED	0	.900	0.000	0.000	0.000	0.000	0.000	0.000	0.000

SYSTEM DESIGN 7- RG,2MAN; C(S) 2.6; C(D) 2.7; C(T) 5.3

ADV.STRATEGY	TACTIC	E(U)	P(SD)	E(N)	E(S)	E(SNM)	P(D)	P(C)	P(ID)
1 ASSAULT	5	.704	.560	1.000	.560	5.600	1.000	.200	1.000
3 DETERRED	0	.700	0.000	0.000	0.000	0.000	0.000	0.000	0.000
4 DETERRED	0	.700	0.000	0.000	0.000	0.000	0.000	0.000	0.000
9 DETERRED	0	.700	0.000	0.000	0.000	0.000	0.000	0.000	0.000
11 DETERRED	0	.900	0.000	0.000	0.000	0.000	0.000	0.000	0.000
12 DETERRED	0	.900	0.000	0.000	0.000	0.000	0.000	0.000	0.000

SYSTEM DESIGN 8- MIN MCA; C(S) 1.1; C(D) 15.4; C(T) 16.5

ADV.STRATEGY	TACTIC	E(U)	P(SD)	E(N)	E(S)	E(SNM)	P(D)	P(C)	P(ID)
1 ASSAULT	5	.704	.560	1.000	.560	5.600	1.000	.200	1.000
3 BREACH,10K	3	.804	.577	1.000	.577	5.766	.200	.099	.100
4 NORMAL,10K	2	.719	.718	2.622	.718	7.181	.532	.280	.283
9 SAMPLER,.3	5	.816	.815	9.143	.815	.245	0.000	.183	.185
11 KILLMCA,10	3	.929	.659	1.000	.659	6.594	.140	.042	.042
12 DETERRED	0	.900	0.000	0.000	0.000	0.000	0.000	0.000	0.000

SYSTEM DESIGN 9- QUANT E; C(S) 3.5; C(D) 5.9; C(T) 9.4

ADV.STRATEGY	TACTIC	E(U)	P(SD)	E(N)	E(S)	E(SNM)	P(D)	P(C)	P(ID)
1 ASSAULT	5	.704	.560	1.000	.560	5.600	1.000	.200	1.000
3 HOSTAGE,10	3	.804	.577	1.000	.577	5.766	.200	.099	.100
4 DETERRED	0	.700	0.000	0.000	0.000	0.000	0.000	0.000	0.000
9 SAMPLER,.3	5	.816	.815	9.143	.815	.245	0.000	.183	.185
11 KILLMCA,10	3	.929	.659	1.000	.659	6.594	.140	.042	.042
12 DETERRED	0	.900	0.000	0.000	0.000	0.000	0.000	0.000	0.000

S,R,M,L OR ABORT ? r

SYSTEM DESIGN 10- PER MON; C(S) 1.7; C(D) 15.4; C(T) 17.1

ADV.STRATEGY	TACTIC	E(U)	P(SD)	E(N)	E(S)	E(SNM)	P(D)	P(C)	P(ID)
1 ASSAULT	5	.704	.560	1.000	.560	5.600	1.000	.200	1.000
3 HOSTAGE,10	3	.804	.577	1.000	.577	5.766	.200	.099	.100
4 NORMAL,10K	2	.719	.718	2.622	.718	7.181	.532	.280	.283
9 SAMPLER,.3	5	.816	.815	9.143	.815	.245	0.000	.183	.185
11 KILLMCA,10	3	.929	.659	1.000	.659	6.594	.140	.042	.042
12 DETERRED	0	.900	0.000	0.000	0.000	0.000	0.000	0.000	0.000

(continued)

TABLE 17. (Concluded)

SYSTEM DESIGN 11- QE,PER ; C(S) 4.1; C(D) 5.9; C(T) 10.0

ADV.STRATEGY	TACTIC	E(U)	P(SD)	E(N)	E(S)	E(SNM)	P(D)	P(C)	P(ID)
1 ASSAULT	5	.704	.560	1.000	.560	5.600	1.000	.200	1.000
3 HOSTAGE,10	3	.804	.577	1.000	.577	5.766	.200	.099	.100
4 DETERRED	0	.700	0.000	0.000	0.000	0.000	0.000	0.000	0.000
9 SAMPLER,.3	5	.816	.815	9.143	.815	.245	0.000	.183	.185
11 KILLMCA,10	3	.929	.659	1.000	.659	6.594	.140	.042	.042
12 DETERRED	0	.900	0.000	0.000	0.000	0.000	0.000	0.000	0.000

SYSTEM DESIGN 12- ALL MCA; C(S) 6.1; C(D) 5.8; C(T) 12.0

ADV.STRATEGY	TACTIC	E(U)	P(SD)	E(N)	E(S)	E(SNM)	P(D)	P(C)	P(ID)
1 ASSAULT	5	.704	.560	1.000	.560	5.600	1.000	.200	1.000
3 HOSTAGE,10	3	.804	.577	1.000	.577	5.766	.200	.099	.100
4 DETERRED	0	.700	0.000	0.000	0.000	0.000	0.000	0.000	0.000
9 DETERRED	0	.700	0.000	0.000	0.000	0.000	0.000	0.000	0.000
11 KILLMCA,10	3	.929	.659	1.000	.659	6.594	.140	.042	.042
12 DETERRED	0	.900	0.000	0.000	0.000	0.000	0.000	0.000	0.000

SYSTEM DESIGN 13- NO ACCT; C(S) 1.5; C(D) 19.0; C(T) 20.5

ADV.STRATEGY	TACTIC	E(U)	P(SD)	E(N)	E(S)	E(SNM)	P(D)	P(C)	P(ID)
1 ASSAULT	5	.704	.560	1.000	.560	5.600	1.000	.200	1.000
3 HOSTAGE,10	3	.804	.577	1.000	.577	5.766	.200	.099	.100
4 NORMAL,.3K	5	1.000	1.000	41.250	33.000	9.900	0.000	0.000	0.000
9 SAMPLER,.3	5	1.000	1.000	10.000	1.000	.300	0.000	0.000	0.000
11 KILLMCA,10	3	.929	.659	1.000	.659	6.594	.140	.042	.042
12 DETERRED	0	.900	0.000	0.000	0.000	0.000	0.000	0.000	0.000

SYSTEM DESIGN 14- ALL ACT; C(S) 6.7; C(D) 15.4; C(T) 22.1

ADV.STRATEGY	TACTIC	E(U)	P(SD)	E(N)	E(S)	E(SNM)	P(D)	P(C)	P(ID)
1 ASSAULT	5	.704	.560	1.000	.560	5.600	1.000	.200	1.000
3 HOSTAGE,10	3	.804	.577	1.000	.577	5.766	.200	.099	.100
4 NORMAL,10K	2	.719	.718	2.622	.718	7.180	.532	.280	.283
9 SAMPLER,.3	5	.789	.788	8.866	.788	.236	0.000	.210	.213
11 KILLMCA,10	3	.929	.659	1.000	.659	6.594	.140	.042	.042
12 DETERRED	0	.900	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Designs 3, 6, and 7 have the best response to Adversary 11 because they all contain roving guards.

- Adversaries 4 and 9 have unlimited, undetected tries in Design 13, which has no accounting system. They can divert all the SNM they desire, totally undetected. This demonstrates one benefit of a nominal accounting system.

Table 18 shows the performance of each system design, without the adversary detail. The performance measures are:

- C_S Safeguards cost ($\$10^6$/year)
- C_D Diversion cost ($\$10^6$/year)
- C_T Total cost ($C_S + C_D$)
- #ATT Frequency of attempt per 1000 years
- #TRY Expected number of tries to steal SNM per attempt
- P_D Probability of ever detecting an adversary given an attempt
- P_{ID} Probability of ever identifying an adversary given an attempt
- P_C Probability of ever capturing an adversary given an attempt
- CSNM Expected quantity (kg) of SNM diverted per year given an attempt
- USNM Unconditional expected quantity (kg) of SNM diverted per 1000 years (or g/y).

TABLE 18. Summary performance measures for all designs.

	DESIGN	C(S)	C(D)	C(T)	#ATT	#TRY	P(D)	P(ID)	P(C)	CSNM	USNM
1	TESTBED	6.65	5.80	12.45	3.0	1.0	.45	.38	.11	6.0	18.0
2	MINIMAL	.50	21.49	21.99	20.0	5.9	.10	.17	.13	4.1	81.1
3	MAXIMAL	12.05	2.66	14.71	1.0	1.0	1.00	1.00	.20	5.6	5.6
4	MEDIUM	1.70	15.41	17.11	17.0	6.6	.23	.24	.19	2.5	42.2
5	2-MAN	2.20	5.80	8.00	3.0	1.0	.45	.38	.11	6.0	18.0
6	R GUARD	2.10	2.75	4.85	12.0	8.5	.08	.25	.18	.7	8.3
7	RG,2MAN	2.60	2.66	5.26	1.0	1.0	1.00	1.00	.20	5.6	5.6
8	MIN MCA	1.10	15.41	16.51	17.0	6.6	.23	.24	.19	2.5	42.2
9	QUANT E	3.55	5.89	9.44	14.0	7.4	.10	.23	.17	1.5	20.7
10	PER MON	1.70	15.41	17.11	17.0	6.6	.23	.24	.19	2.5	42.2
11	QE,PER	4.15	5.89	10.04	14.0	7.4	.10	.23	.17	1.5	20.7
12	ALL MCA	6.15	5.80	11.95	3.0	1.0	.45	.38	.11	6.0	18.0
13	NO ACCT	1.50	19.02	20.52	17.0	13.9	.08	.07	.02	3.0	51.0
14	ALL ACT	6.70	15.41	22.11	17.0	6.4	.23	.25	.21	2.5	42.1

Design Sensitivity

Fourteen designs are evaluated in Table 18. Designs 1, 2, 3, and 4 span the range of system designs; the rest are sensitivity studies. These four are:

- 1 Testbed design (our approximation)
- 2 Minimal safeguards
- 3 Maximal safeguards
- 4 Moderate safeguards (less than testbed).

Of the four, the testbed design is best on a total cost basis. Adding the last few safeguards to get the maximal system (Design 3) costs $5.4 million/year, but reduces diversion costs by only $2.28 million/year.

While these four designs contain the lowest and highest diversion cost C_D and safeguards cost C_S, they do not contain the best overall system (based on C_T). This was found during the sensitivity analysis of Designs 5 through 14.

Design 6 is best overall, according to the C_T criterion. It depends more heavily on physical security than on MC&A systems, using only process monitors and personnel monitors, and it has a nominal accounting system to detect repeated attempts. It also costs 68% less than the testbed and its diversion costs are 53% lower. Notice, however, that it does not rank first on either P_D or P_C minimization criteria.

Marginal Component Benefit

The information in Tables 17 and 18 can be used to evaluate the incremental benefit of individual safeguards components as well as safeguards subsystems. A sample comparison is shown in Table 19. There are enough sensitivity cases in Table 18 to evaluate most components. For instance, the first row in Table 19 shows the net change in costs when the quantity estimation component (1.1) is added to a similar design without quantity estimation. (Design 8 is the base case; Design 9 has quantity estimation equipment.) Safeguards costs go up by $\$2.45 \cdot 10^6$/y, and diversion costs drop by $\$9.5 \cdot 10^6$/y. This is a net decrease in total cost, so for this base case the quantity estimation equipment is well worth the price.

TABLE 19. Marginal evaluation safeguards components.

Subsystem component	Marginal impact ΔC_S		Marginal value ΔC_D
1. Material control			
1.1 Quantity estimation	+2.45		-9.50
1.2 Process state	--	Required MC	--
1.3 Personnel monitors	+0.6		0
1.4 Procedure monitors	+2.0		-0.11
2. Physical security			
2.1 Stationary guards	--	Required PS	--
2.2 Roving guards	+0.4		-3.14
2.3 Two-person rule	+0.5		-0.09
3.3 Accounting			
3.1 Nominal system	+0.2		-3.59
3.2 Frequent inventory	+5.0		0.0

Personnel monitors add nothing to facility security. According to Table 18: detection doesn't matter for outsiders; Adversary 3 takes the plant manager hostage; Adversary 4 steals SNM during working hours when he or she is supposed to be near SNM; and Adversaries 11 and 12 disable all monitors. The only adversary for whom personnel monitors are in the diversion path is Adversary 9, and he or she gets less than 0.03 kg/y. Thus, personnel monitors do nothing to decrease total costs; therefore, they are not worth the investment for this illustrative data set.

In contrast, increasing physical security is beneficial. We assume that stationary guards are required. Adding the two-person rule is marginally useful; adding roving guards is even more valuable.

The nominal records system has value because it stops repeating adversaries. Frequent inventories are assumed to be very expensive, and since much of the SNM is diverted by adversaries who only try once, they are not worth the improved information they provide.

PERFORMANCE GRAPHS

Figure 22 plots the information contained in the first three columns and P_D column in Table 18 for selected system designs. The horizontal axis is P_D, and the vertical axis plots are C_S, C_D, and C_T. Designs 5, 6, 7, 8, 9, and 13 appear in Fig. 22, although they are not all plotted on each cost curve. This is because the only designs plotted on each curve are those with minimum cost for that curve. The lower curve in Fig. 22 is the minimum safeguards cost C_S for the set of designs included on the graph. The middle curve is diversion cost, and the upper curve is total cost C_T.

Using the criterion of minimum total cost, the best design is Design 6. This system has a probability of detection equal to 0.18. A probability of detection equal to 1.0 could be obtained by spending only slightly more and using Design 7 rather than Design 6. However, this added detection performance has little value, since diversion cost does not decrease much for Design 3.

Figure 23 graphs costs versus another performance measure--the expected amount of SNM diverted per year. The results here are similar to the previous case. Design 6 is the minimum point on the total cost curve. Design 7, with slightly less SNM diverted, costs $\$9.5 \cdot 10^6$ more. Note that the direction of the curve is reversed relative to Fig. 22: costs rise greatly as more SNM is diverted.

We can also plot the alternative performance measures:

- P_{ID} Probability of identification
- P_C Probability of capture
- #ATT The expected number of attempts per 1000 years.

The probability plots will resemble Fig. 22. The shape of the curve is reversed for #ATT, as in Fig. 23.

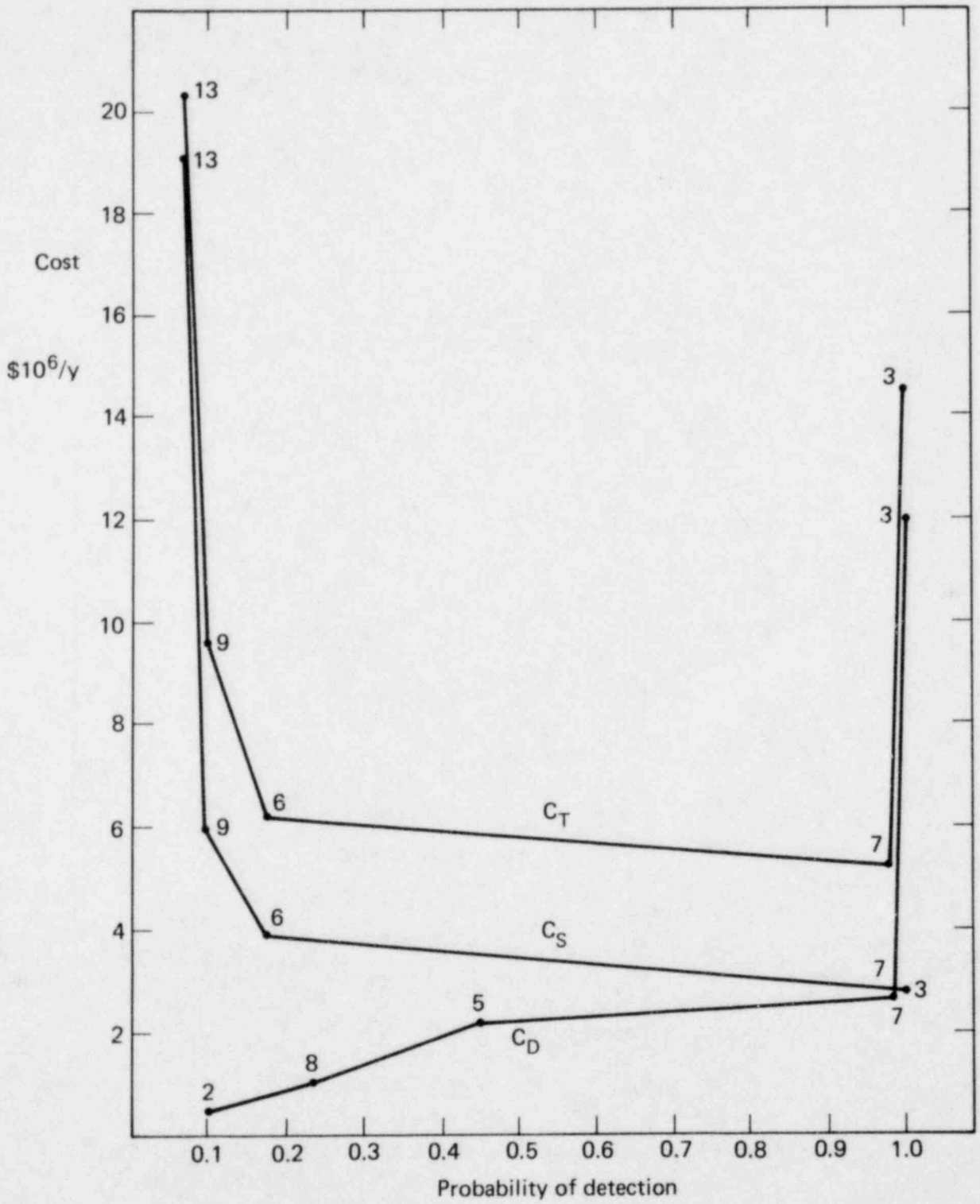


FIG. 22. Performance graph: cost versus probability of detection.

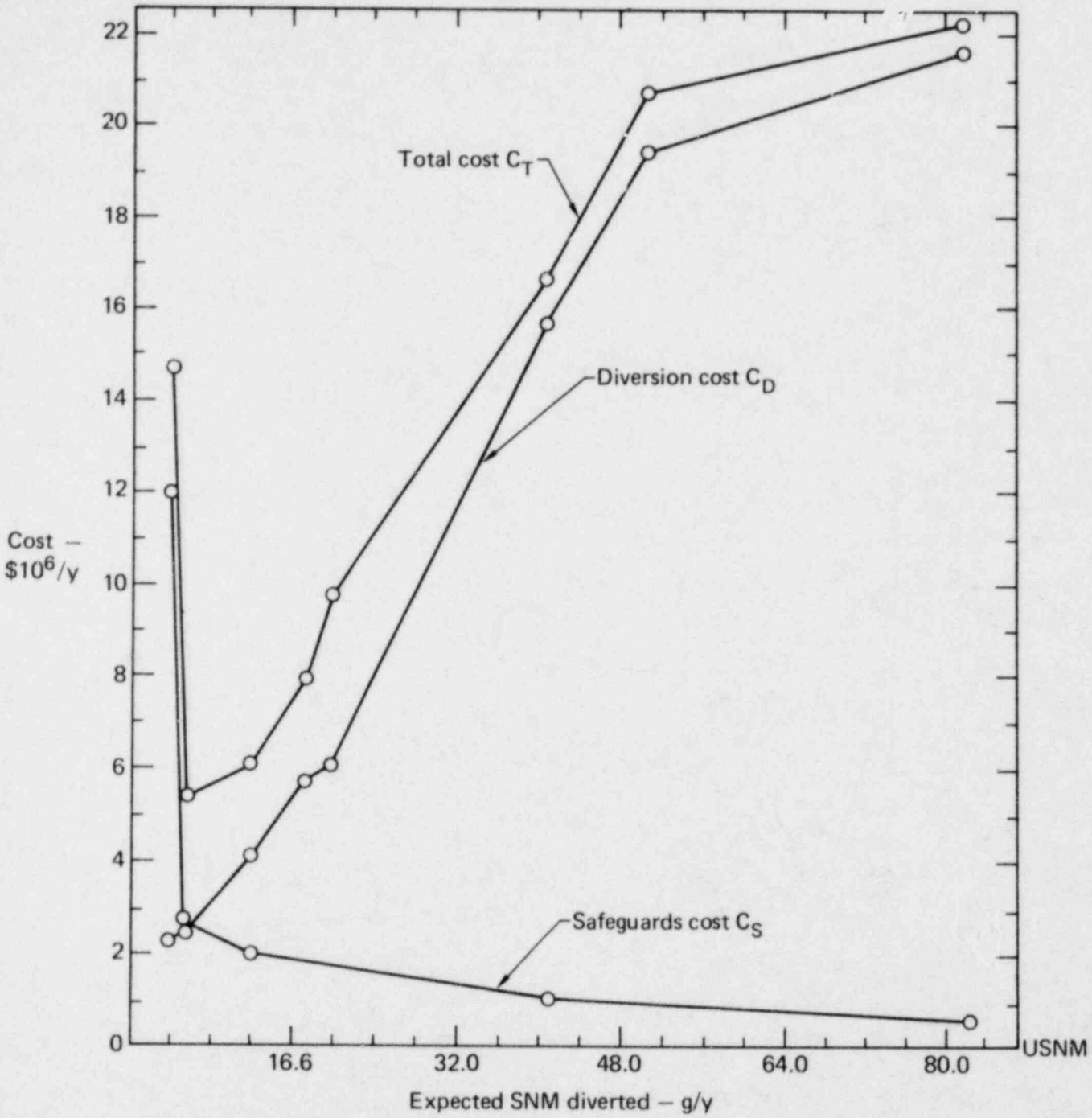


FIG. 23. Performance graph: cost versus expected SNM diverted per y.

Figure 24 shows C_D on the horizontal axis, and C_S on the vertical axis. All 14 designs are plotted. This curve shows the tradeoff between safeguards cost and diversion cost. The only designs on the efficient frontier are 2, 6, 7, and 8. As the relative weighting between C_S and C_D changed in computing C_T (it is 1 to 1 in this report), a different design becomes optimal. The slope of the straight line in Fig. 24 is -1. The point of tangency between the curve and the line gives the optimal system--Design 7 in this case.

Figure 24 also demonstrates the sensitivity of optimal design to C_D. If C_D is weighted more than 38% of C_S, then Design 7 will be optimal for the designs depicted. Notice that a weighting of 56% on C_D would be required before Design 3 would be optimal. This graph shows how the optimal design would change when numbers are varied in the Consequence Model. If the weighting between C_S and C_D is kept at 1 to 1, then C_D could be reduced by 91% or increased by a factor of 5.6, and still Design 6 would be optimal. Of course, the robustness of Design 6 most likely would be less if there were more than 14 designs plotted in Fig. 24.

The implication of this result is that the design decision is virtually insensitive to an increase in consequences. A doubling of the number of deaths and injuries resulting from a nuclear detonation in the Consequence Model would result in an 81% increase in the expected consequences of a 5-kg diversion; C_D for all designs increases by slightly less than 80%. This still does not make Design 3 nearly as attractive as Design 6.

ADVERSARY SENSITIVITY

A sensitivity case that is likely to have impact on the design decision is to vary P_1--the utility of "no try"--in the Adversary Utility Model. This, in effect, changes the adversary's attitude toward capture, and hence the choice of strategy and tactics. It also changes his or her willingness to make the attempt, which would influence the optimal level of plant safeguards if the criteria were based on adversary utility. Tables 20 and 21 show the results of a model run with these assumptions:

- P_1 = 0.1 for all adversaries ("fearless" adversaries)
- P_2 and P_3 remain unchanged.

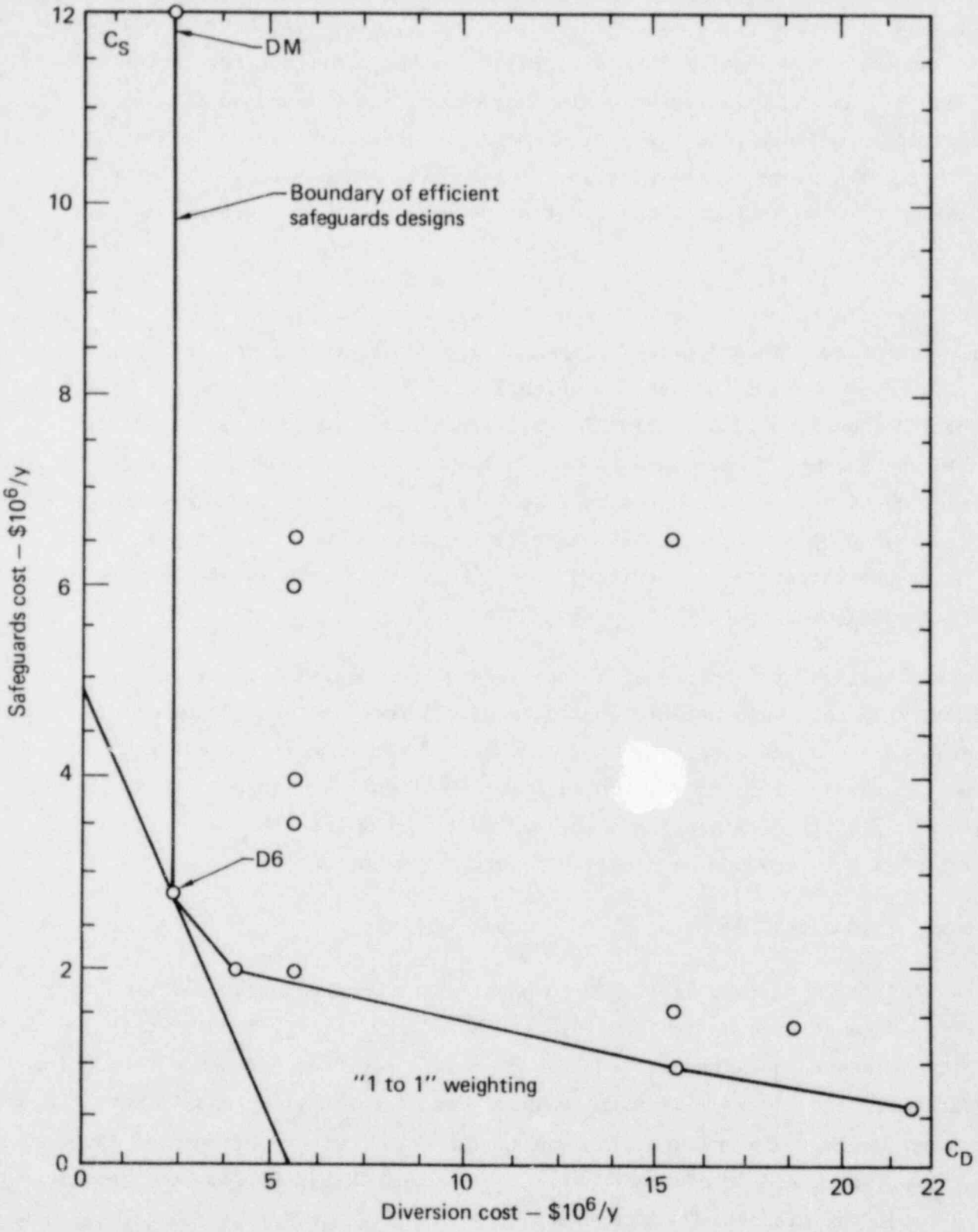


FIG. 24. Value-impact tradeoff curve: safeguards cost versus diversion cost.

TABLE 20. Adversary decisions--"fearless" adversary sensitivity.

SYSTEM DESIGN 1; C(S) 6.6; C(D) 12.7; C(T) 19.3

ADV.STRAT	TACTIC	E(U)	P(SD)	E(N)	E(S)	E(SNM)	P(D)	P(C)	P(ID)
1.1	5	.584	.560	1.000	.560	5.600	1.000	.200	1.000
3.4	5	.690	.673	1.000	.673	6.733	.200	.158	.160
4.3	5	.311	.311	1.379	.311	3.113	.976	.687	.694
9.0	0	.100	0.000	0.000	0.000	0.000	0.000	0.000	0.000
11.1	5	.763	.745	1.000	.745	7.446	.140	.069	.070
12.3	5	.837	.837	2.036	.837	8.371	.250	.163	.164

SYSTEM DESIGN 2; C(S) .5; C(D) 21.5; C(T) 22.0

ADV.STRAT	TACTIC	E(U)	P(SD)	E(N)	E(S)	E(SNM)	P(D)	P(C)	P(ID)
1.1	5	.584	.560	1.000	.560	5.600	1.000	.200	1.000
3.1	5	.820	.800	1.000	.800	8.000	0.000	0.000	0.000
4.3	3	.894	.894	2.305	.894	8.942	.150	.105	.107
9.1	5	.815	.815	9.143	.815	.245	0.000	.183	.185
11.2	5	.820	.800	1.000	.800	8.000	0.000	0.000	0.000
12.4	5	.998	.998	2.238	.998	9.980	0.000	.002	.002

SYSTEM DESIGN 3; C(S) 12.0; C(D) 10.6; C(T) 22.7

ADV.STRAT	TACTIC	E(U)	P(SD)	E(N)	E(S)	E(SNM)	P(D)	P(C)	P(ID)
1.1	5	.584	.560	1.000	.560	5.600	1.000	.200	1.000
3.4	5	.456	.445	1.000	.445	4.452	.560	.444	.448
4.3	5	.279	.279	1.339	.279	2.791	.980	.719	.727
9.0	0	.100	0.000	0.000	0.000	0.000	0.000	0.000	0.000
11.1	5	.606	.591	1.000	.591	5.913	.527	.261	.264
12.3	5	.647	.646	1.798	.646	6.465	.587	.353	.356

SYSTEM DESIGN 4; C(S) 1.7; C(D) 18.2; C(T) 19.9

ADV.STRAT	TACTIC	E(U)	P(SD)	E(N)	E(S)	E(SNM)	P(D)	P(C)	P(ID)
1.1	5	.584	.560	1.000	.560	5.600	1.000	.200	1.000
3.4	5	.690	.673	1.000	.673	6.733	.200	.158	.160
4.3	2	.718	.718	2.622	.718	7.181	.532	.280	.283
9.1	5	.815	.815	9.143	.815	.245	0.000	.183	.185
11.1	5	.763	.745	1.000	.745	7.446	.140	.069	.070
12.3	3	.863	.863	2.429	.863	8.632	.250	.136	.138

○ Denotes changed decision from Figure 4.9

Case: P_1=.1 for all adversaries -- change from Figure 3.12

TABLE 21. Summary performance measures--"fearless" adversary sensitivity.

DESIGN	C(S)	C(D)	C(T)	#ATT	#TRY	P(D)	P(ID)	P(C)	CSNM	USNM
1	6.65	12.68	19.33	9.0	1.5	.63	.42	.33	6.03	54.2
2	.50	21.47	21.97	20.0	5.9	.10	.17	.13	4.05	81.1
3	12.05	10.62	22.67	9.0	1.4	.83	.55	.46	4.86	43.7
4	1.70	18.21	19.91	20.0	5.9	.27	.23	.18	3.50	69.9
5	2.20	15.55	17.75	20.0	2.1	.82	.64	.59	3.08	61.5
6	2.10	15.62	17.72	20.0	5.8	.36	.29	.25	2.89	57.9
7	2.60	13.39	15.99	20.0	1.9	.85	.70	.65	2.52	50.5
8	1.10	18.21	19.31	20.0	5.9	.27	.23	.18	3.50	69.9
9	3.55	15.89	19.44	20.0	5.8	.28	.25	.21	3.23	64.6
10	1.70	18.21	19.91	20.0	5.9	.27	.23	.18	3.50	69.9
11	4.15	15.89	20.04	20.0	5.8	.28	.25	.21	3.23	64.6
12	6.15	14.44	20.59	20.0	1.9	.83	.65	.61	2.95	59.0
13	1.50	21.81	23.31	20.0	12.2	.14	.08	.04	3.94	78.8
14	6.70	18.20	24.90	20.0	5.8	.27	.24	.20	3.48	69.6

Table 20 shows the strategy and tactics each adversary will use under the first four system designs. This table can be compared with Table 17, in which the circles indicate where decisions have been changed relative to Table 17. With one exception, the circled numbers in the tactics column are all T5's. As the adversary values SNM more and fears capture less, he or she will abort fewer tries. Threfore the adversaries will be choosing Tactic T5, "never abort."

Table 21 can be compared with Table 18. These are the results for each system design if the adversary assigns a utility value of 0.1 to the status quo. Because more adversaries will attempt the diversion, the expected SNM diverted and the diversion cost both rise. Even so, Design 7 remains optimal, showing that system design is relatively insensitive to adversary utility.

SUMMARY

We have shown several illustrative results and insights produced by this type of analysis. These illustrations show how the model output changes because of varying input parameters. In addition, they demonstrate how various performance measures might be used in safeguards criteria setting.

IV. CONCLUSIONS

This report documents the dynamic version of the Aggregated Systems Model and demonstrates its performance. The ASM has two primary advantages over earlier models: (1) it represents more accurately the adversary choices among strategies and tactics, and (2) it reflects the dynamic nature of repeated attempts. An assumption central to the model is that probabilities of detection, identification, and capture do not change over time; that is, neither the adversary nor the system gets smarter--with the exception of late detection or identification.

The analytic process has generated a potentially important and useful measure of deterrence. The adversary is deterred when the expected utility derived from the diversion attempt is less than the utility of not trying. The deterrence measure depends upon adversary preferences and safeguards performance.

In this report, we have shown the use of various performance measures, in addition to expected adversary utility; these include P_D, P_{ID}, P_C, and E(SNM). The type of system being evaluated determines the choice among these measures: for example, an MC&A system evaluation might depend on P_D or E(SNM); a physical security evaluation might depend on P_{ID} and P_C.

REFERENCES

1. C. A. Bennett, W. M. Murphy, and T. S. Sherr, Societal Risk Approach to Safeguards Design and Evaluation, U.S. Energy Research and Development Administration, Washington, D.C., ERDA-7 (June 1975).
2. I. Sacks, et al., Material Control System Design: Test Bed Nitrate Storage Areas, Lawrence Livermore Laboratory, Livermore, Calif. (May 1978).
3. R. A. Howard, Dynamic Probabilistic Systems, Volumes 1 & 2, Wiley & Sons, New York (1971).
4. C. W. Kirkwood and S. M. Pollack, Methodology for Characterizing Potential Adversaries of Nuclear Safeguards Systems, Woodward-Clyde Consultants, San Francisco, Calif. (November 1978).
5. D. Kaul, et al., Consequences of Adversary Actions in the Nuclear Power Fuel Cycle, Science Applications, Inc., Chicago, forthcoming.
6. Methodology and Preliminary Models for Analyzing Nuclear Safeguards Decisions, Applied Decision Analysis, Inc., Menlo Park, Calif. (November 1978).

NMS/n11

APPENDIX A:
ADDITIONAL INFORMATION ON ADVERSARY TACTICS

For each diversion strategy, the number of tactics that an adversary must consider is a function of the number of subsystems that might give timely detection. Given n subsystems, there are 2^n possible combinations of subsystems that might detect the attempt. We call this combination of possible detections by subsystems a "detection condition." For each of the 2^n detection conditions, the adversary must decide whether or not to abort. Under one condition, the decision is obvious: clearly, an adversary would not abort if he or she were not detected. Thus, there are $2^n-1 = y$ detection conditions for which an adversary must make an abort/no abort decision.

We define a tactic as a rule that says "abort" or "no abort" for each of these y possible conditions. There are 2^y different ways to say "abort" or "no abort" for the y conditions. Therefore, there are 2^y or $2^{(2^n-1)}$ tactics for n subsystems. For three subsystems, there are 128 tactics to evaluate.

A drawing, such as Fig. A-1, is useful for identifying these possible tactics. This figure shows the conditions--events where a subsystem or combination of subsystems does or does not detect--and the adversary's decision on whether or not to abort under that condition. This tree illustrates a specific example with three subsystems. There are 2^3, or eight, detection conditions, all but one of which is followed by an abort/no abort decision. This yields the 15 endpoints on the tree. (In general, there are $2^{n+1}-1$ endpoints.)

A tactic can be specified by the set of endpoints that the adversary could reach by following the abort/no abort decisions for each condition. If, for example, he or she chooses the tactic "always abort if detected by any subsystem," he or she could end up at endpoint 1 (detection by SS1), endpoint 9 (detection by SS2), endpoint 13 (detection by SS3), or endpoint 15 (no detection). If detection by Subsystem 3 did not increase the adversary

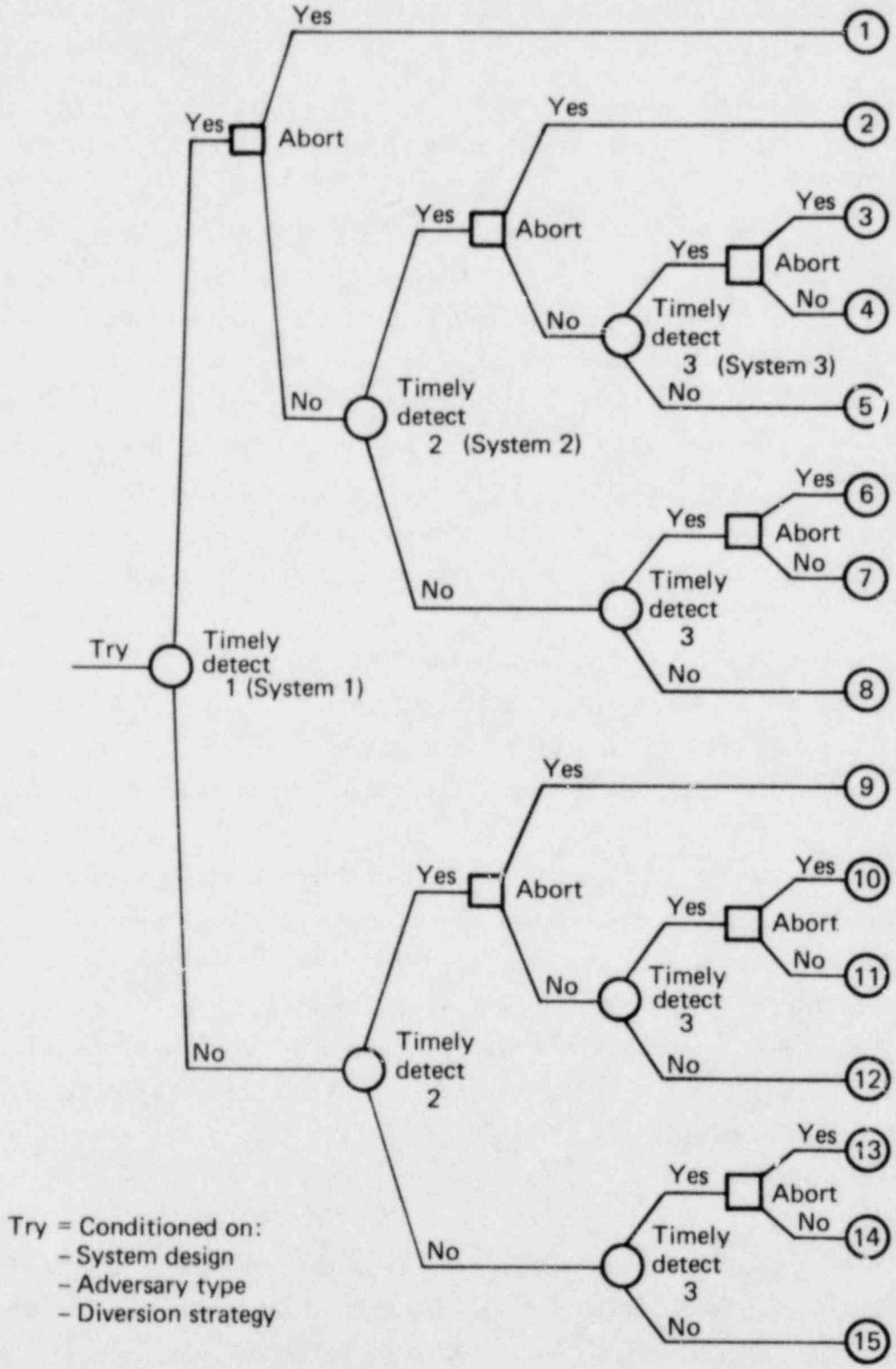


FIG. A-1. Diversion tactics analysis.

probability of identification and capture, the adversary would not abort if only SS3 detected, yielding the set of end points: 1, 9, 14, and 15. Similarly, endpoints 1, 10, 13, and 15 are possible if the adversary continues when only SS2 detects.

Obviously, the endpoints for a tactic do not include both the "abort" and "no abort" branches emanating from a particular detection condition. For example, an adversary cannot reach both endpoints 3 and 4, which follow detection by all three subsystems, because one indicates aborting and the other continuing. Endpoint 15, on the other hand, is possible for every tactic. Of the other 14 endpoints, only 7 can be chosen for any 1 tactic. Every branch that says "abort" has a corresponding branch that says "no abort,"* and only one endpoint from each pair can be chosen. Therefore, there are $2^7 = 128$ possible tactics in this case.

Many of these 128 tactics are not logically consistent if detections by subsystems are independent. For example, a rational adversary probably would not choose to abort an attempt if detected by Subsystem 1 alone, but would choose "no abort" if detected by both 1 and 2. Examination of the 128 tactics shows that all but 19 of them are logically inconsistent. All but these 19 can be eliminated from further consideration. A table of all logically consistent tactics is included in Table A-1, which lists these 19 possible tactics.

In the illustrative data set utilized in the model, only Subsystems 1 and 2 could detect in a timely fashion; Subsystem 3 (accounting) could only detect a diversion at a later time, after a completed diversion. This reduced the number of feasible tactics to five:

1. (1, 9, 15) Always abort if detected by any subsystem
2. (1, 12, 15) Abort if detected by Subsystem 1, but continue otherwise
3. (2, 8, 9, 15) Abort if detected by Subsystem 1, but continue otherwise

*Pairs are 1&8, 2&5, 3&4, 6&7, 9&12, 10&11, 13&14.

TABLE A-1. Possible endpoints in Figure A-1.

Tactic	Set of possible endpoints in Figure A-1	
1.	1, 9, 13, 15	- Always abort
2.	1, 9, 14, 15	
3.	1, 10, 12, 13, 15	
4.	1, 10, 12, 14, 15	
5.	1, 11, 12, 14, 15	
6.	2, 6, 8, 9, 13, 15	
7.	2, 6, 8, 9, 14, 15	
8.	2, 6, 8, 10, 12, 13, 15	
9.	2, 6, 8, 10, 12, 14, 15	
10.	2, 6, 8, 11, 12, 14, 15	
11.	2, 7, 8, 9, 14, 15	
12.	2, 7, 8, 10, 12, 14, 15	
13.	2, 7, 8, 11, 12, 14, 15	
14.	3, 5, 6, 8, 10, 12, 13, 15	
15.	3, 5, 6, 8, 10, 12, 14, 15	
16.	3, 5, 6, 8, 11, 12, 14, 15	
17.	3, 5, 7, 8, 10, 12, 14, 15	
18.	3, 5, 7, 8, 11, 12, 14, 15	
19.	4, 5, 7, 8, 11, 12, 14, 15	- Never abort

4. (2, 8, 12, 15) Abort if detected by both Subsystems 1 and 2, but continue otherwise
5. (5, 8, 12, 15) Never abort.

For each tactic, a decision tree similar to the one in Fig. 7 could be constructed.

Notice that the tree in Fig. 7 represents only the adversary's perceptions. In actuality, additional branches, not shown, could reflect the possibility of late detection; obviously, at the time of the diversion, the adversary cannot know whether he or she will be detected later. However, we assume that the adversary does know (before he or she begins the attempt) the probabilities of timely and late detection, as well as the resulting probabilities of identification associated with each.

Since the adversary does not know whether or not late detection will occur, the best that he or she can do is to make the tactical decision after considering explicitly the possibility of late detection, identification, and interruption in addition to the possibility of never being detected. Figure A-2 shows a portion of Fig. 7 representing the actual probability tree (for one subsystem) and the adversary's observable tree. It also shows the adversary's calculation of probabilities for the observable tree.

The Markov model discussed in Section I considers only the simple case in which no late identification occurs. In reality, late identification is extremely important: it allows the safeguards authority eventually to identify and stop the adversary from making repeated attempts. Therefore, late identification must be included in the Markov model. Figure A-3 shows what happens to the Markov process when late identification is added. We define "x" as the number of tries between the try on which the late identification process is initiated and the try on which the identification is made. The variable x is calculated as the frequency of inspections that might detect an adversary late, divided by the frequency of adversary tries.

The probabilities of being captured or identified remain the same as they were in the case with no late identification. Every repeatable try has some probability of never being detected, k_2; in this case, the adversary remains in the Markov process at the top of Fig. A-3. But there is some probability, k_1, that a repeatable try will lead to late identification. If an adversary does something to cause late identification, he or she will have x more tries before being identified and captured or identified but escaping capture--unless he or she stops because all desired material is acquired. In either case, after x more tries, the adversary drops into the process illustrated at the

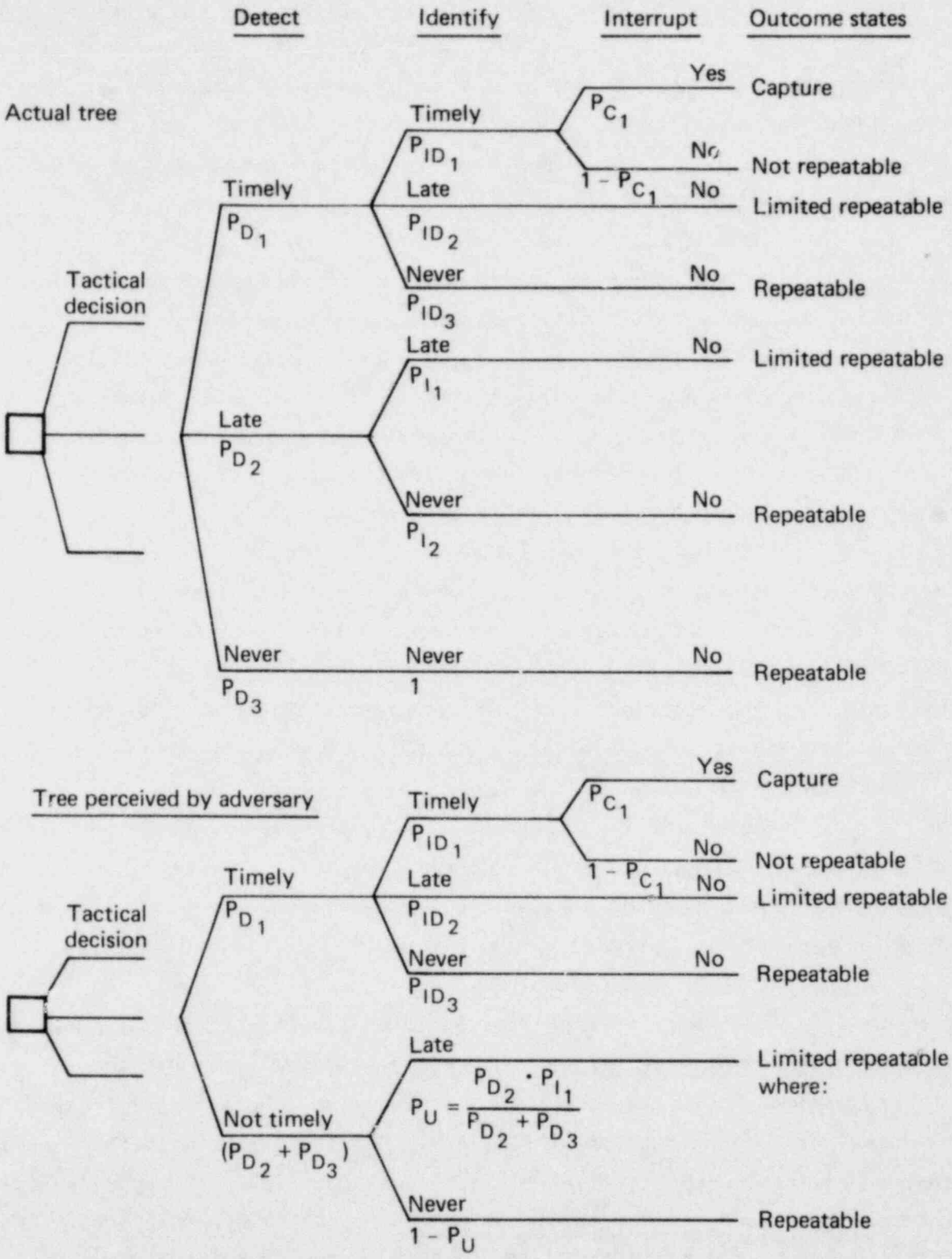


FIG. A-2. Actual and perceived probability trees.

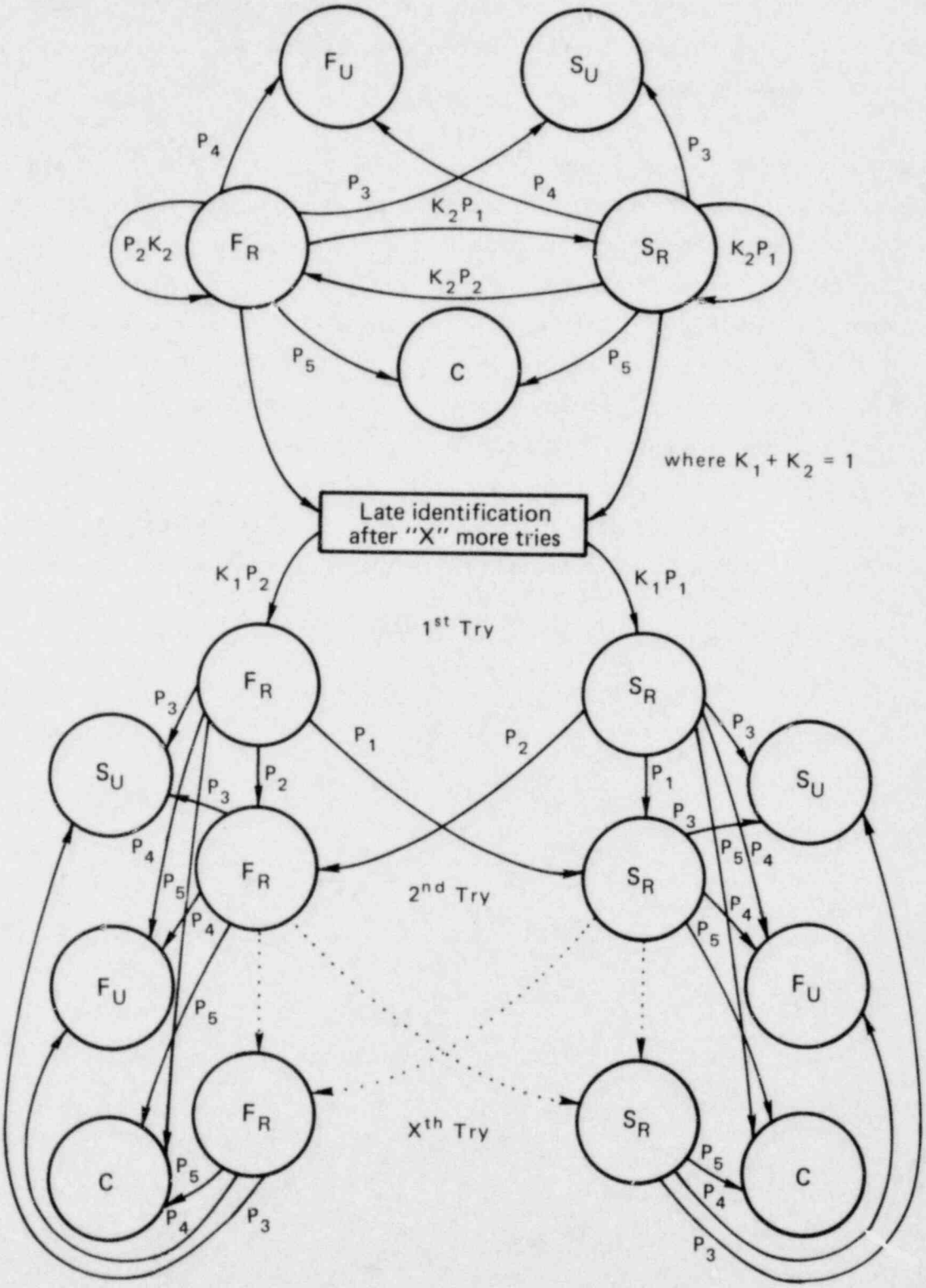


FIG. A-3. Markov model with late identification.

bottom of Fig. A-3 once the late identification process is triggered. The next x-1 tries appear exactly the same as a try in the regular Markov process. It is on the x^{th} try that probabilities change, sending the adversary to one of the three nonrepeatable states.

The calculation of long-run probabilities for the outcome lottery becomes a two-step process. For the first x tries, the process behaves as it did in the simple Markov model. But on the x+1 try and every following try, the probability of late identification (brought about by the try x time units before) must be considered. Each try before the x+1 try has a probability of $(p_1 + p_2)$ that it will be repeatable; that is, no identification has occurred. At each try after x, the probability of "not being identified" drops from $(p_1 + p_2)$ to $k_2(p_1 + p_2)$ and the probabilities of "capture" and "unrepeatable but not captured" states rise in proportion to k_1, where k_1 is the portion of repeatable tries that will be identified late. Actual calculations of probabilities and performance measures follow in Appendix B.

APPENDIX B:

MATHEMATICAL DERIVATION OF RESULTS

The derivation of these algorithms relies heavily on the use of Markov statistics and geometric transformations. Transforming a discrete function, f(n), into a geometric summation, $f_g(z) = \sum_{n=0}^{\infty} f(n)z^n$, follows from the observation that the series expansion of $f_g(z)$ has unique coefficients f(n). Because the expansion is unique, the relationship between the discrete function and the transform is also unique. For a more detailed discussion of geometric transforms (also known as z-transforms) consult texts such as Dynamic Probabilistic Systems by Ronald Howard.[3]

First calculate the probability that exactly m units of SNM are stolen, given identification at try n.

$$p(s = m|n) = \binom{n-1}{m} P_s^m P_F^{n-m-1} P_C + \binom{n}{m} P_s^m P_F^{n-m} P_I$$

The first term reflects the probability that m units were diverted on the first n-1 tries and capture occurred on the nth trial. Note: Capture implies that no SNM was successfully diverted on that trial.

The second term reflects the probability that m units were diverted on n trials and identification but not capture occurred on the nth try. The adversary might or might not have diverted SNM on the nth trial; both are possible.

P_S = Probability of success on a try.

$P_F = 1-P_S$ = Probability of failure on a try.

P_C = Probability of capture given identification.

P_I = Probability of escape given identification.

$P_R = 1-P_C-P_I$ = Probability of no identification, allowing the attempt to be repeated.

The probability that n trials occurred, with identification occurring on the nth trial is:

$$P(n) = \begin{cases} P_R^{n-1}\left(1-P_R\right) & \text{if } n \leq x \\ P_R^{n-1}\, k_2^{n-x-1}\left(1-P_R\, k_2\right) & \text{if } n \geq x+1 \end{cases}$$

where x = number of trials until late identification can occur.

k_1 = proportion of repeatable attempts that will be identified late.

k_2 = proportion of repeatable attempts that will never be identified; $k_1 + k_2 = 1.0$.

The probability that m units are diverted is:

$$P(s=m) = \sum_{n=m}^{x} p\left(s=m|n\right) \cdot P_R^{n-1}\left(1-P_R\right) \qquad \text{Note that } \binom{n}{m} = 0 \text{ if } n < m$$

$$+ \sum_{n=x+1}^{\infty} P(s=m|n) \cdot P_R^{n-1}\, k_2^{n-x-1}\left(1-P_R\, k_2\right)$$

$$= \sum_{n=m+1}^{x} \binom{n-1}{m} P_S^m\, P_F^{n-m-1}\, P_C\, P_R^{n-1}\left(1-P_R\right) + \sum_{n=m}^{x} \binom{n}{m} P_S^m\, P_F^{n-m}\, P_I\, P_R^{n-1}\left(1-P_R\right)$$

$$+ \sum_{n=x+1}^{\infty} \binom{n-1}{m} P_S^m\, P_F^{n-m-1}\, P_C\, P_R^{n-1}\, k_2^{n-x-1}\left(1-k_2\, P_R\right) + \sum_{n=x+1}^{\infty} \binom{n}{m} P_S^m\, P_F^{n-m}\, P_I$$

$$P_R^{n-1}\, k_2^{n-x-1}\left(1-k_2\, P_R\right)$$

Note that $\sum_{n=a}^{\infty} = \sum_{n=0}^{\infty} - \sum_{n=0}^{a-1}$

$$\sum_{n=x+1}^{\infty} = \sum_{n=m+1}^{\infty} \binom{n-1}{m} P_S^m \; P_F^{n-m-1} \; P_C \; P_R^{n-1} \; k_2^{n-x-1} \left(1-k_2 \; P_R\right) - \sum_{n=m+1}^{x} \binom{n-1}{m}$$

$$P_S^m \; P_F^{n-m-1} P_C \; P_R^{n-1} \; k_2^{n-x-1} \left(1-k_2 \; P_R\right)$$

$$\sum_{n=x+1}^{\infty} = \sum_{n=m}^{\infty} \binom{n}{m} P_S^m \; P_F^{n-m} \; P_I \; P_R^{n-1} \; k_2^{n-x-1} \left(1-k_2 \; P_R\right) - \sum_{n=m}^{x} \binom{n}{m} P_S^m \; P_F^{n-m}$$

$$P_I \; P_R^{n-1} \; k_2^{n-x-1} \left(1-P_R \; k_2\right)$$

Using geometric transforms:

$$\sum_{n=m+1}^{\infty} \binom{n-1}{m} P_S^m \; P_F^{n-m-1} \; P_C \; P_R^{n-1} \; k_2^{n-x-1} \left(1-k_2 \; P_R\right) = P_S^m \; P_R^m \; P_C \; k_2^{m-x}$$

$$\left(1-k_2 \; P_R\right) \sum_{r=0}^{\infty} \binom{r+m}{m} \left(P_F \; P_R \; k_2\right)^r, \text{ where } r=n-m-1$$

$$= \frac{P_S^m \; P_R^m \; P_C \; k_2^{m-x} \left(1-k_2 \; P_R\right)}{\left(1-P_F \; P_R \; k_2\right)^{m+1}}$$

$$\sum_{n=m}^{\infty} \binom{n}{m} P_S^m \; P_F^{n-m} \; P_I \; P_R^{n-1} \; k_2^{n-x-1} \left(1-k_2 \; P_R\right) = P_S^m \; P_R^{m-1} \; P_I \; k_2^{m-x-1}$$

$$\left(1-k_2 \; P_R\right) \sum_{j=0}^{\infty} \binom{j+m}{m} \left(P_F \; P_R \; k_2\right)^j, \text{ where } j=n-m$$

$$= \begin{cases} \dfrac{P_S^m \; P_I \; k_2^{m-x-1} \; P_R^{m-1} \left(1-k_2 \; P_R\right)}{\left(1-P_F \; P_R \; k_2\right)^{m+1}} & \text{if } m \neq \emptyset \\ \dfrac{P_I \left(1-k_2 \; P_R\right)}{\left(1-P_F \; P_R \; k_2\right) k_2^{x+1} P_R} - \dfrac{P_I \left(1-k_2 \; P_R\right)}{k_2^{x+1} \; P_R} = \dfrac{P_I \; P_F \left(1-k_2 \; P_R\right)}{\left(1-P_F \; P_R \; k_2\right) k_2^x} & \text{if } m = \emptyset \end{cases}$$

$$\text{Then } P(S=m) = \frac{P_S^m \; P_R^m \; P_C \; k_2^{m-x} \left(1-k_2 \; P_R\right)}{\left(1-P_F \; P_R \; k_2\right)^{m+1}} + \frac{P_S^m \; P_I \; k_2^{m-x-1} \; P_R \left(1-k_2 \; P_R\right)}{\left(1-P_F \; P_R \; k_2\right)^{m+1}}$$

$$- \sum_{n=m+1}^{x} \binom{n-1}{m} P_S^m \; P_F^{n-m-1} \; P_C \; P_R^{n-1} \; k_2^{n-x-1} \left(1-P_R \; k_2\right)$$

$$- \sum_{n=m}^{x} \binom{n}{m} P_S^m \; P_F^{n-m} \; P_I \; P_R^{n-1} \; k_2^{n-x-1} \left(1-k_2 \; P_R\right)$$

$$+ \sum_{n=m+1}^{x} \binom{n-1}{m} P_S^m \; P_F^{n-m-1} \; P_C \; P_R^{n-1} \left(1-P_R\right)$$

$$+ \sum_{n=m}^{x} \binom{n}{m} P_S^m \; P_F^{n-m} \; P_I \; P_R^{n-1} \left(1-P_R\right)$$

$$P(S=m) = \begin{cases} \dfrac{P_S^m \; P_R^{m-1} \; k_2^{m-x-1} \left(1-k_2 \; P_R\right) \left(P_C \; P_R \; k_2 + P_I\right)}{\left(1-P_F \; P_R \; k_2\right)^{m+1}} \\ + \displaystyle\sum_{n=m+1}^{x} \binom{n-1}{m} P_S^m \; P_F^{n-m-1} \; P_C \; P_R^{n-1} \left(1-P_R - k_2^{n-x-1} \left(1-k_2 \; P_R\right)\right) \\ + \displaystyle\sum_{n=m}^{x} \binom{n}{m} P_S^m \; P_F^{n-m} \; P_I \; P_R^{n-1} \left(1-P_R - k_2^{n-x-1} \left(1-k_2 \; P_R\right)\right) \quad \text{if } m \neq \emptyset \\ \dfrac{\left(1-k_2 \; P_R\right)\left(P_C + P_I \; P_F\right)}{\left(1-P_F \; P_R \; k_2\right) \; k_2^x} + \left(P_F \; P_I + P_C\right) \displaystyle\sum_{n=1}^{x} P_F^{n-1} \; P_R^{n-1} \left(1-P_R - k_2^{n-x-1} \left(1-k_2 \; P_R\right)\right) \\ \text{if } m = \emptyset \end{cases}$$

The probability that the adversary will divert all the SNM he or she wants is: $P(m \geq S_D)$, where S_D x unit/attempt is the quantity he or she desires. If the adversary diverts S_D units, he or she will be satisfied and discontinue the attempt.

$$P\left(m \geq S_D\right) = \sum_{m=S_D}^{\infty} P(m) = \sum_{m=S_D}^{\infty} \frac{P_S^m\ P_R^{m-1}\ k_2^{m-x-1}\ T_1\left(1-k_2\ P_R\right)}{\left(1-P_F\ P_R\ k_2\right)^{m+1}} \left(\text{Where } T_1 = P_C\ P_R\ k_2 + P_I\right)$$

$$+ \sum_{m=S_D}^{\infty} \left(\sum_{n=m+1}^{x} \binom{n-1}{m} P_S^m\ P_F^{n-m-1}\ P_C\ P_R^{n-1} \left(1-P_R-k_2^{n-x-1}\ \left(1-k_2\ P_R\right)\right)\right.$$

$$\left. + \sum_{n=m}^{x} \binom{n}{m}\ P_S^m\ P_F^{n-m}\ P_I\ P_R^{n-1}\ \left(1-P_R-\ k_2^{n-x-1}\ \left(1-k_2\ P_R\right)\right)\right)$$

Using Transforms

$$\sum_{m=S_D}^{\infty} \frac{P_S^m\ P_R^{m-1}\ k_2^{m-x-1}\ T_1\left(1-P_R\ k_2\right)}{\left(1-P_F\ P_R\ k_2\right)^{m+1}} = \sum_{r=o}^{\infty} \frac{P_S^{S_D+r}\ T_1\ P_R^{S_D+r-1}\ k_2^{S_D+r-x-1}\left(1-P_R\ k_2\right)}{\left(1-P_F\ P_R\ k_2\right)^{S_D+r+1}}$$

$$= \frac{P_S^{S_D}\ P_R^{S_D-1}\ k_2^{S_D-x-1}\ T_1\left(1-P_R\ k_2\right)}{\left(1-P_F\ P_R\ k_2\right)^{S_D+1}} \sum_{r=o}^{\infty} \left(\frac{P_S\ P_R\ k_2}{1-P_F\ P_r\ k_2}\right)^r \quad \text{where } r=m-S_D$$

$$= \frac{T_1\ P_s^{S_D}\ P_R^{S_D-1}\ k_2^{S_D-x-1}\left(1-k_2\ P_R\right)}{\left(1-P_F\ P_R\ k_2\right)^{S_D+1}} \left(\frac{1}{1-\left(\frac{P_S\ P_R\ k_2}{1-P_F\ P_R\ k_2}\right)}\right)$$

$$= \frac{T_1\ P_S^{S_D}\ P_R^{S_D-1}\ k_2^{S_D-x-1}\left(1-P_R\ k_2\right)}{\left(i-P_F\ P_R\ k_2\right)^{S_D+1}} \frac{\left(1-P_F\ P_R\ k_2\right)}{\left(1-P_R\ k_2\right)} = \frac{P_S^{S_D}\ P_R^{S_D-1}\ k_2^{S_D-x-1}\ T_1}{\left(1-P_F\ P_R\ k_2\right)^{S_D}}$$

Then, given that $S_D \neq 0$,

$$P\left(m \geq S_D\right) = \frac{P_S^{S_D}\ P_R^{S_D-1}\ k_2^{S_D-x-1}\ T_1}{\left(1-P_F\ P_R\ k_2\right)^{S_D}} + \sum_{m=S_D}^{x} \sum_{n=m+1}^{x} \binom{n-1}{m} P_S^m\ P_F^{n-m-1}\ P_C\ P_R^{n-1}$$

$$\bullet \left(1-P_R-k_2^{n-x-1}\ \left(1-k_2\ -P_R\right)\right)$$

$$+ \sum_{m=S_D}^{x} \sum_{n=m}^{x} \binom{n}{m} P_S^m P_F^{n-m} P_I P_R^{n-1} \left(1-P_R - k_2^{n-x-1} \left(1-k_2 P_R\right)\right)$$

$P(m \geq 0) = 1$

Note: Because the limits on the inner sum are m and x where m is determined by the outer sum, the outer sum is bounded above by x also.

Expected SNM

Next, the expected amount of SNM diverted is calculated.

$$E(m) = \sum_{m=0}^{S_D-1} m \; P(S=m) + \sum_{m=S_D}^{\infty} S_D \; P(S=m)$$

$$= \sum_{m=0}^{S_D-1} m \; \frac{P_S^m P_R^{m-1} k_2^{m-x-1} T_1 \left(1-P_R k_2\right)}{\left(1-P_F P_R - k_2\right)^{m+1}}$$

$$+ \sum_{m=0}^{S_D-1} m \sum_{n=m+1}^{x} \binom{n-1}{m} P_S^m P_F^{n-m-1} P_C P_R^{n-1} \left(1-P_R - k_2^{n-x-1} \left(1-k_2 P_R\right)\right)$$

$$+ \sum_{m=0}^{S_D-1} m \sum_{n=m}^{x} \binom{n}{m} P_S^m P_F^{n-m-1} P_I P_R^{n-1} \left(1-P_R - k_2^{n-x-1} \left(1-k_2 P_R\right)\right)$$

$$+ \sum_{m=S_D}^{\infty} S_D \; \frac{P_S^m P_R^{m-1} k_2^{m-x-1} T_1 \left(1-P_R k_2\right)}{\left(1-P_F P_R k_2\right)^{m+1}}$$

$$+ \sum_{m=S_D}^{x} S_D \sum_{n=m+1}^{x} \binom{n-1}{m} P_S^m P_F^{n-m-1} P_C P_R^{n-1} \left(1-P_R - k_2^{n-x-1} \left(1-k_2 P_R\right)\right)$$

$$+ \sum_{m=S_D}^{x} S_D \sum_{n=m}^{x} \binom{n}{m} P_S^m P_F^{n-m} P_I P_R^{n-1} \left(1-P_R - k_2^{n-x-1} \left(1-k_2 P_R\right)\right)$$

Evaluating the first single sum:

$$\sum_{m=0}^{S_D-1} m \frac{P_S^m P_R^{m-1} k_2^{m-x-1} T_1\left(1-P_R k_2\right)}{\left(1-P_F P_R k_2\right)^{m+1}} = \sum_{m=0}^{\infty} m \frac{P_S^m P_R^{m-1} k_2^{m-x-1} T_1\left(1-P_R k_2\right)}{\left(1-P_F P_R k_2\right)^{m+1}}$$

$$- \sum_{m=S_D}^{\infty} m \frac{P_S^m P_R^{m-1} k_2^{m-x-1} T_1\left(1-P_R k_2\right)}{\left(1-P_F P_R k_2\right)^{m+1}}$$

$$\sum_{m=0}^{\infty} m \frac{P_S^m P_R^{m-1} k_2^{m-x-1} T_1\left(1-P_R k_2\right)}{\left(1-P_F P_R k_2\right)^{m+1}} = \frac{T_1\left(1-P_R k_2\right)}{\left(1-P_F P_R k_2\right) P_R k_2^{x+1}} \sum_{m=0}^{\infty} m \left(\frac{P_S P_R k_2}{1-P_F P_R k_2}\right)^m$$

$$= \frac{T_1\left(1-P_R k_2\right)}{\left(1-P_F P_R k_2\right) P_R k_2^{x+1}} \frac{\left(\frac{P_S P_R k_2}{1-P_F P_R k_2}\right)}{\left(1-\frac{P_S P_R k_2}{1-P_F P_R k_2}\right)^2} = \frac{T_1 P_S}{k_2^x \left(1-P_S k_2\right)}$$

$$\sum_{m=S_D}^{\infty} m \frac{P_S^m P_R^{m-1} k_2^{m-x-1} T_1\left(1-P_R k_2\right)}{\left(1-P_F P_R k_2\right)^{m+1}} = \sum_{r=0}^{\infty} \binom{r+S_D}{} \frac{P_S^{r+S_D} P_R^{r+S_D-1} k_2^{r+S_D-x-1} T_1\left(1-P_R k_2\right)}{\left(1-P_F P_R k_2\right)^{r+S_D+1}}$$

Where $m=r+S_D$

$$= \sum_{r=0}^{\infty} r \frac{P_S^{S_D+r} P_R^{S_D-1+r} k_2^{S_D-x-1+r} T_1\left(1-P_R k_2\right)}{\left(1-P_F P_R k_2\right)^{r+S_D+1}} + \sum_{m=S_D}^{\infty} S_D \frac{P_S^m P_R^{m-1} k_2^{m-x-1} T_1\left(1-P_R k_2\right)}{\left(1-P_F P_R k_2\right)^{m+1}}$$

$$\sum_{r=0}^{\infty} r \frac{P_S^{S_D+r} P_R^{S_D+r-1} k_2^{S_D+r-x-1} T_1\left(1-P_R k_2\right)}{\left(1-P_F P_R k_2\right)^{r+S_D+1}} = \left(\frac{P_S P_R k_2}{1-P_F P_R k_2}\right)^{S_D}.$$

$$\sum_{r=0}^{\infty} r \frac{P_S^r P_R^{r-1} k_2^{r-x-1} T_1\left(1-P_R k_2\right)}{\left(1-P_F P_R k_2\right)^{r+1}} = \left(\frac{P_S P_R k_2}{1-P_F P_R k_2}\right)^{S_D} \frac{T_1 P_S}{k_2^x \left(1-P_S k_2\right)}$$

Then

$$\sum_{m=0}^{S_D-1} m \frac{P_S^m P_R^{m-1} k_2^{m-x-1} T_1 \left(1-P_R k_2\right)}{\left(1-P_F P_R k_2\right)^{m+1}} = \frac{T_1 P_S}{k_2^x \left(1-P_R k_2\right)} \left(1 - \left(\frac{P_S P_R k_2}{1-P_F P_R k_2}\right)^{S_D}\right)$$

$$- \sum_{m=S_D}^{*\infty} S_D \frac{P_S^m P_R^{m-1} k_2^{m-x-1} T_1 \left(1-P_R k_2\right)}{\left(1-P_F P_R k_2\right)^{m+1}}$$

*Note that this last summation is the inverse of the second summation in the original equation and they will cancel.

Putting all the pieces together:

$$E(m) = \frac{T_1 P_S}{k_2^x \left(1-P_S k_2\right)} \left(1 - \left(\frac{P_S P_R k_2}{1-P_F P_R k_2}\right)^{S_D}\right)$$

$$+ \sum_{m=0}^{S_D-1} m \sum_{n=m+1}^{x} \binom{n-1}{m} P_S^m P_C P_F^{n-m-1} P_R^{n-1} \left(1-P_C - \left(k_2^{n-x-1} \left(1-k_2 P_R\right)\right)\right)$$

$$+ \sum_{m=0}^{S_D-1} m \sum_{n=m}^{x} \binom{n}{m} P_S^m P_R^{n-1} P_I P_F^{n-m} \left(1-P_C-k_2^{n-x-1} \left(1-k_2 P_R\right)\right)$$

$$+ \sum_{m=S_D}^{x} S_D \sum_{n=m+1}^{x} \binom{n-1}{m} P_S^m P_F^{n-m-1} P_C P_R^{n-1} \left(1-P_C-k_2^{n-x-1} \left(1-k_2 P_R\right)\right)$$

$$+ \sum_{m=S_D}^{x} S_D \sum_{n=m}^{x} \binom{n}{m} P_S^m P_F^{n-m} P_I P_R^{n-1} \left(1-P_C-k_2^{n-x-1} \left(1-k_2 P_R\right)\right)$$

Expected Utility

$$E(UTIL) = P(S \geq S_D) \cdot U(S_D) + U(C) \cdot \left(\sum_{m=0}^{S_D-1} P(S=m \text{ and } C) \right)$$

$$+ \left(\sum_{m=1}^{S_D-1} U(m)\ p(S=m \text{ and } I) \right) + U(NONE) \cdot P(S=0 \text{ and } I)$$

Where $U(S_D) = 1$ = the utility of acquiring the desired SNM.

$U(m)$ = the utility of acquiring m units of SNM.

$U(NONE)$ = the utility of having no SNM but not being captured.

$U(C) = 0$ = the utility of being captured.

P(Capture) = probability of capture given identification.

I = adversary was identified but escaped.

$$\sum_{m=1}^{S_D-1} U(m)\ p(S=m \text{ and } I) = \sum_{m=1}^{S_D-1} am \cdot p(S=m \text{ and } I) + \sum_{m=1}^{S_D-1} b \cdot p(s=m \text{ and } I)$$

$$\sum_{m=1}^{S_D-1} am\ \frac{P_S^m\ P_R^{m-1}\ P_I\ k_2^{m-x-1}\left(1-k_2\ P_R\right)}{\left(1-P_F\ P_R\ k_2\right)^{m+1}} = \frac{a\ P_I\ P_S}{k_2^x\left(1-P_R\ k_2\right)}\left(1-\left(\frac{P_S\ P_R\ k_2}{1-P_F P_R k_2}\right)^{S_D}\right)$$

$$- \sum_{m=S_D}^{\infty} a\ \frac{S_D\ P_S^m\ P_R^{m-1}\ k_2^{m-x-1}\left(1-P_R\ k_2\right)}{\left(1-P_F\ P_R\ k_2\right)^{m+1}}$$

(See derivation on page 99)

$$\sum_{m=1}^{S_D-1} b\ \frac{P_S^m\ P_R^{m-1}\ P_I\ k_2^{m-x-1}\left(1-k_2\ P_R\right)}{\left(1-P_F\ P_R\ k_2\right)^{m+1}} = \sum_{m=0}^{\infty} b\ \frac{P_S^m\ P_R^{m-1}\ P_I\ k_2^{m-x-1}\left(1-k_2\ P_R\right)}{\left(1-P_F\ P_R\ k_2\right)^{m+1}}$$

$$- \sum_{m=S_D}^{\infty} b\ \frac{P_S^m\ P_R^{m-1}\ P_I\ k_2^{m-x-1}\left(1-k_2\ P_R\right)}{\left(1-P_F\ P_R\ k_2\right)^{m+1}} - \frac{b\ P_I\left(1-k_2\ P_R\right)}{P_R\ k^{x+1}\left(1-P_F\ P_R\ k_2\right)}$$

$$= \frac{b\ P_I}{P_R\ k_2^{x+1}} \left(1-\left(\frac{P_S\ P_R\ k_2}{1-P_F P_R k_2}\right)^{S_D}\right) - \frac{b\ P_S \left(1-k_2\ P_R\right)}{P_R\ k_2^{x+1}\left(1-P_F\ P_R\ k_2\right)}$$

Similarly

$$\sum_{m=S_D}^{\infty} \frac{a\ S_D\ P_S^m\ P_R^{m-1}\ k_2^{m-x-1}\ P_I\left(1-P_R\ k_2\right)}{\left(1-P_F\ P_R\ k_2\right)^{m+1}} = \frac{a\ S_D\ P_S^{S_D}\ P_R^{S_D-1}\ k_2^{S_D-x-1}\ P_I}{\left(1-P_F\ P_R\ k_2\right)^{S_D}}$$

Then

$$E\ (UTIL) = P\ (S \geq S_D) + \frac{a\ P_I\ P_S}{k_2^x\left(1-P_R\ k_2\right)}\left(1-\left(\frac{P_S\ P_R\ k_2}{1-P_F\ P_R\ k_2}\right)^{S_D} - \frac{\left(P_S\ P_R\ k_2\right)^{S_D-1}\ S_D\left(1-P_R\ k_2\right)}{\left(1-P_F\ P_R\ k_2\right)^{S_D}}\right)$$

$$+ \frac{b\ P_I}{P_R\ k_2^{x+1}}\left(1-\left(\frac{P_S\ P_R\ k_2}{1-P_F\ P_R\ k_2}\right)^{S_D}\right) - \frac{b\ P_I\left(1-k_2\ P_R\right)}{P_R\ k_2^{x+1}\left(1-P_F\ P_R\ k_2\right)}$$

$$+ U\ (NONE) \cdot \frac{P_I\ P_F\left(1-k_2\ P_R\right)}{\left(1-P_F\ P_R\ k_2\right)\ k_2^x} + U\ (NONE) \sum_{n=1}^{x} P_F^n\ P_R^{n-1}\ P_I\left(1-P_R-k_2^{n-x-1}\left(1-k_2\ P_R\right)\right)$$

$$+ \sum_{m=1}^{S_D-1} am+b \sum_{n=m}^{x}\binom{n}{m} P_S^m\ P_F^{n-m}\ P_R^{n-1}\ P_I\left(1-P_R\ k_2^{n-x-1}\left(1-P_R\ k_2\right)\right)$$

The probability of eventually being identified is calculated exactly like the probability of success with the exception that no success occurs if the adversary is identified but escapes.

$$P\ (S_L=m) \sum_{n=m+1}^{x} \binom{n-1}{m} P_S^m\ P_F^{n-x-1}\ P_R^{n-1}\left(1-P_R\right) + \sum_{n=x+1}^{\infty} \binom{n-1}{m} P_S^m\ P_F^{n-m-1} P_R^{n-1} k_2^{n-x-1}\left(1-k_2 P_R\right)$$

$$= \frac{P_R^m\ P_S^m\ k_2^{m-x}\left(1-k_2\ P_R\right)}{\left(1-P_F\ P_R\ k_2\right)^{m+1}} + \sum_{n=m+1}^{x} \binom{n-1}{m} P_S^m\ P_F^{n-m-1}\ P_R^{n-1}\left(1-P_R-k_2^{n-x-1}\left(1-k_2\ P_R\right)\right)$$

(See calculations on page 94 and 95)

$$P\ (m \geq S_D) = \sum_{m=S_D}^{\infty} P\ (S_L=m)$$

$$= \sum_{m=S_D}^{\infty} \frac{P_R^m\ P_S^m\ k_2^{m-x}\left(1-k_2\ P_R\right)}{\left(1-P_F\ P_R\ k_2\right)^{m+1}} + \sum_{m=S_D}^{\infty} \sum_{n=m+1}^{x} \binom{n-1}{m} P_S^m\ P_F^{n-m-1} P_R^{n-1}\left(1-P_R-k_2^{n-x-1}\left(1-k_2 P_R\right)\right)$$

$$= \frac{P_S^{S_D}\ P_R^{S_D}\ k_2^{S_D-x}}{\left(1-P_F\ P_R\ k_2\right)^{S_D}} + \sum_{m=S_D}^{x} \sum_{n=m+1}^{x} \binom{n-1}{m} P_S^m\ P_F^{n-m-1}\ P_R^{n-1}\left(1-P_R-k_2^{n-x-1}\left(1-k_2\ P_R\right)\right)$$

(See calculations on page 98)

$P\ (TRAP) = 1 - P\ (m \geq S_D)$, i.e., if the adversary does not steal his or her desired amount and quit, the adversary will be identified - trapped - unable to repeat attempts.

GLOSSARY

This glossary is organized by the following major categories: Adversary Model, Facility Safeguards Model, Adversary-Facility Interaction, Attempt Outcomes, and Performance Measures. Within categories, the organization is usually chronological according to attempt sequence of events. The following list shows the major categories and the terms defined within each category:

ADVERSARY MODEL
- Adversary
- SNM
- Attempt
- Try
- Strategy
- Diversion path
- Monitor target set
- Tactic
- Detection condition

FACILITY SAFEGUARDS MODEL
- System design
- Subsystem
- Component
- Definitions of all subsystems and components in the illustrative data set

ADVERSARY-FACILITY INTERACTION
- Detection
- Identification
- Interruption
- Timing
- Acquire SNM
- Divert SNM

ATTEMPT OUTCOMES
- Capture
- Success
- Failure
- Repeatable attempt
- Not repeatable attempt

PERFORMANCE MEASURES
- Safeguards cost
- Diversion cost
- Total cost
- Expected annual attempts
- Expected tries per attempt

PERFORMANCE MEASURES (Cont'd)
- Probability of detection given attempt
- Probability of identification and capture given attempt
- Expected adversary utility

Expected adversary tries	Deterrence
Unconditional expected annual diversion	Prevention
Conditional expected annual diversion	Response

ADVERSARY MODEL

Adversary: People who might attempt to divert (steal) special nuclear material (SNM). They are generally classified according to their group size, employee status, equipment resources, special knowledge or authority, and desired quantity of SNM. Earlier analyses referred to the adversaries in terms of "adversary action sequences" or "sequence categories."

SNM (Special Nuclear Material): Plutonium or highly enriched uranium from which an explosive or dispersal weapon could be made.

Attempt: The initiation of an adversary's plan to divert SNM. An attempt implies both the existence of an adversary and the initiation of his or her plan to divert. If the adversary exists but is deterred from initiating the plan, then no attempt has occurred. An attempt may be composed of several tries, and it may last as long as several weeks or months.

Try: An adversary's action to acquire and divert some quantity of SNM. Generally, a try lasts for less than one day.

Strategy (number of tries, quantity per try, diversion path): The elements of an adversary's plan: (1) the number of tries in the attempt, (2) the quantity to be acquired on each try, and (3) the diversion path. One alternative that can be considered by an adversary is the "no attempt" strategy.

Diversion Path: The set of safeguards the adversary encounters on his path to and from the SNM. Equivalent to MTS.

Monitor Target Sets (MTS): The set of safeguards components the adversary will encounter during the try.

Tactic: A decision rule that governs the adversary's reaction to detection by the safeguards system. Responses are: "abort the try" and "no abort."

Detection Condition: The set of safeguards subsystems that have detected a particular adversary's try.

FACILITY SAFEGUARDS MODEL

System Design: A collection of safeguards components (aggregated into subsystems) installed in a facility.

Subsystem: A collection of similar safeguards components that have been grouped together for two reasons: (1) Adversary tactics (abort/no abort) depend on which safeguards subsystems (not individual components) detect him or her, and (2) the probabilities of adversary identification (by the safeguards authority) depend on which subsystem--not component--detects the adversary.

Component: The elemental unit of the safeguards system in the ASM. In reality, ASM components are composed of many monitors, guards, and accounting records. However, in the ASM, these safeguards elements are called "components" in the aggregate.

Subsystem 1 (SS1)--Electronic Detection: The set of electronic safeguards components. SS1 detection usually results in an electronic signal being sent to a control panel. This system does not identify the adversary without additional information. Detection can be timely, late, or never.

Component 1.1 (Quantity Estimators): Bubblers and other monitors that measure changes in liquid volume.

Component 1.2 (Process State): Instruments that measure pressure, temperature, and other physical characteristics of fluids in process. Also, sensors that give valve positions, fluid flow, etc.

Component 1.3 (Personnel Monitors): Floormats, infrared area monitors, and other instruments that detect the presence of individuals in an area.

Component 1.4 (Procedure Monitors): Computers that monitor personnel access, valve position changes, and other actions by operators.

Subsystem 2 (SS2)--Visual Detection: Guards and other human safeguards elements. SS2 detection usually results from observing the

adversary or observing that something is amiss physically. This system usually identifies the adversary. Detection can be timely, late, or never.

Component 2.1 (Guard Stations): Stationary guards, usually found at the plant gate or storage vaults.

Component 2.2 (Roving Guards): Self explanatory.

Component 2.3 (Two-Person Rule): Two operators present for all operations.

Subsystem 3 (SS3)--Accounting: Records and procedures to track the plant's SNM inventory. SS3 detection results from a discrepancy in the books. This system does not identify the adversary without additional information. Detection can be only late or never.

Component 3.1 (Nominal Accounting System): The accounting procedures required by the NRC.

Component 3.2 (Frequent Inventory): Assay of all SNM in the plant. This extensive and expensive procedure requires plant shutdown for one to two months to flush all pipes and vats, shake down filter bags, etc.

Inspection Frequency: The minimum time interval between the adversary's try and the first possible late identification. This term applies to a complete system design, and not to individual subsystems or components. Inspections reveal discrepancies in records as well as physical security violations.

Safeguards Authority: The person, group, or computer with the authority to order intervention by the security force in order to interrupt the adversary's try. The safeguards authority is able to identify the adversary if given sufficient information.

ADVERSARY-FACILITY INTERACTION

Detection: The safeguards authority receives a signal (electronic or verbal) that one or more components have sensed a predefined diversion condition. The detection signal may be a false alarm.

Identification: The safeguards authority determines that an attempt is occurring or has occurred; i.e., detection was not a false alarm. In addition, the safeguards authority knows who the physical security force should confront in order to interrupt the diversion.

Interruption: The physical security force stops the diversion and prevents any more SNM from crossing the plant boundary. We assume that this includes capturing the adversary and preventing malevolent use of any SNM diverted in the total attempt.

Timing for Detection or Identification (Timely, Late, Never):

Timely -- Physical security has sufficient time to interrupt the try on which detection occurred.

Late -- Detection or identification occurs after the try is complete, but perhaps in time to interrupt a future try.

Never -- The safeguards authority does not detect or identify the adversary on a particular try.

Acquire SNM: Being physically able to take possession of a quantity of SNM during a try. This does not necessarily mean leaving the plant with the SNM.

Divert SNM: The act of leaving the plant with a quantity of SNM.

ATTEMPT OUTCOMES

Capture: The adversary is apprehended and cannot use SNM for malevolent purposes.

Success: The adversary is not captured and diverts some SNM during the try.

Failure: The adversary is not captured, but fails to acquire SNM on a particular try.

Repeatable Attempt (Limited, Unlimited):

Unlimited: The adversary may try again indefinitely.

Limited: The adversary is identified late and has a limited number of tries before identification occurs.

Not Repeatable Attempt: The adversary is identified and will be captured if he or she shows up again at the plant.

NRC FORM 335 (7-77)

U.S. NUCLEAR REGULATORY COMMISSION

BIBLIOGRAPHIC DATA SHEET

1. REPORT NUMBER *(Assigned by DDC)*
NUREG CR/1140 Vol. II
UCRL 52712 Vol. II

4. TITLE AND SUBTITLE *(Add Volume No., if appropriate)*
Aggregated Systems Model of Nuclear Safeguards, Vol. II

2. *(Leave blank)*

3. RECIPIENT'S ACCESSION NO.

7. AUTHOR(S)

5. DATE REPORT COMPLETED
MONTH: November | YEAR: 1979

9. PERFORMING ORGANIZATION NAME AND MAILING ADDRESS *(Include Zip Code)*
NRC Safeguards Program, L-97
Lawrence Livermore Laboratory
P. O. Box 808
Livermore, CA 94550

DATE REPORT ISSUED
MONTH: February | YEAR: 1980

6. *(Leave blank)*

8. *(Leave blank)*

12. SPONSORING ORGANIZATION NAME AND MAILING ADDRESS *(Include Zip Code)*
Dr. Robert L. Shepard
Technical Support Branch
Division of Safeguards, Fuel Cycle & Environmental Research
Office of Nuclear Regulatory Research
Washington, DC 20555

10. PROJECT/TASK/WORK UNIT NO.

11. CONTRACT NO.

13. TYPE OF REPORT
NUREG

PERIOD COVERED *(Inclusive dates)*

15. SUPPLEMENTARY NOTES

14. *(Leave blank)*

16. ABSTRACT *(200 words or less)*

When setting the performance criteria for systems that safeguard special nuclear material (SNM), decision makers must consider characteristics of the adversaries who attempt to divert SNM, safeguards responses to these attempts, costs of safeguards systems, and the consequences of diverted SNM.

This report describes an Aggregated Systems Model that is designed to assist decision makers in integrating and evaluating these diverse factors consistently. The report summarizes the results obtained from applying the procedures required, substitution of electronics for human safeguards, and overall performance criteria for safeguards systems. New performance criteria designed to measure how safeguards systems deter adversary attempts are also described.

17. KEY WORDS AND DOCUMENT ANALYSIS

17a. DESCRIPTORS

17b. IDENTIFIERS/OPEN-ENDED TERMS

18. AVAILABILITY STATEMENT
Unlimited

19. SECURITY CLASS *(This report)*
Unclassified

21. NO. OF PAGES
120

20. SECURITY CLASS *(This page)*
Unclassified

22. PRICE
$