LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

LLNL-JRNL-554944

Exploitation of Ambiguous Cues to Infer Terrorist Activity

K. Ni, D. Faissol, T. Edmunds, R. Wheeler

May 3, 2012

Decision Analysis

Disclaimer

This document was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government nor Lawrence Livermore National Security, LLC, nor any of their employees makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or Lawrence Livermore National Security, LLC. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or Lawrence Livermore National Security, LLC, and shall not be used for advertising or product endorsement purposes.

Submitted to *Decision Analysis*
manuscript (Please, provide the mansucript number!)

Exploitation of Ambiguous Cues to Infer Terrorist Activity

Kevin Ni, Daniel Faissol, Thomas Edmunds, and Richard Wheeler
Lawerence Livermore National Laboratory, {ni1,faissol1,edmunds2,wheeler21}@llnl.gov

To aid intelligence analysts in processing ambiguous data regarding nuclear terrorism threats, we develop a methodology that captures and accounts for the uncertainty in new information and incorporates prior beliefs on likely nuclear terrorist activity. This methodology can guide the analyst when making difficult decisions regarding what data is most critical to examine and what threats require greater attention. To do this, we first characterize the general process of a nuclear terrorist attack on the US. We give a hypothetical attack scenario and describe how this can be characterized on an event tree. Then we define hypothetical cues for this attack scenario and give notional strengths to each cue. Our methodology is based on a Bayesian statistical approach to incorporate ambiguous cues to update prior beliefs of adversary activity. We also perform sensitivity analysis on how cue strengths can affect inference. The method can be used to help support decisions regarding resource allocation and interdiction.

Key words: Bayesian analysis; ambiguous information; nuclear counter-terrorism national security; event tree; decision support; decision analysis

1. Introduction

Intelligence analysts must process large volumes of diverse and noisy data in order to identify indicators that an adversary is planning or executing a nuclear terrorist attack. The objective of this paper is to describe a methodology that exploits indicators, or cues, to better infer adversary strategies and/or activities to guide the analyst when making difficult decisions regarding what data is more critical to examine and what threats require greater attention. We show how ambiguous cues can be used to update prior assessments and support decisions on focusing or deploying additional data collection or interdiction resources. Cues could take many forms, some examples include: adversary surveillance of potential targets prior to an attack, purchase of a boat or plane capable of transporting a weapon to target, or communication between potential adversaries and individuals with nuclear weapons expertise.

Previous security researchers have sketched out taxonomies of ambiguous cues. Chengara (2004) presents one such unrestricted example; the authors classify the types of indicators of terrorist activities into three categories, "Red Flags," "Probable Indicators," and "Possible Indicators," giving examples for each category. A "red flag" indicator could be the reported theft of explosives from a construction site, and a "possible" indicator could be a sudden burst or decline in communications. In this paper, we focus on how to utilize a combination of weak or ambiguous indicators to draw an inference on overall adversary strategies and activities. Several law enforcement affiliated efforts have also detailed possible indicators of terrorist activity (Wyllie (2011), SCN (2005)).

Additionally, Khalsa (2004) provides some examples of indicators of terrorist activity and a process by which these may be qualitatively combined in a tool to determine terrorist activity. However, this work lacks a solid analytical foundation to highlight combinations of ambiguous cues.

Krupka (2006) using Bayes' Rule to update the perceived probability of attack using the probability of a warning given an attack, the authors perform an analysis on how ambiguous information affects decision making processes. Smith (2005) discusses the detectability of adversary actions throughout terrorist activity. With our methodology, we use the combined power of many ambiguous cues to detect adversary activity.

While there are notable general treatments, there are few studies focused on the process of using associated cues to analyze nuclear terrorism risks from non-state actors. Related published research, mostly qualitative, on evidence of attack planning has been done for terrorism in general.

One study on combining weak indicators was done by Smith et al. (2006). The authors studied the spatial and temporal patterns of past terrorist activities using planning and preparatory criminal activity indicators such as gun show attendance, vehicle thefts, parking violations, and phone calls.

Bayesian networks and hidden Markov models have been employed in the analysis of terrorist activity though often without indicators or warnings (Tu et al. (2004)). Our work extends this research by matching a set of observed or non-observed cues to an ensemble of possible threat scenarios and by imposing a rigorous mathematical framework for analysis. Bier (2007) and Hausken and Zhuang (2011) are examples of research that approach modeling terrorist strategy in a game theoretic manner.

Our approach integrates three key elements to produce an updated estimate of adversary behavior in the face of new data: cue observations, a prior estimate of adversary behavior, and an updating methodology. *Cue observations* come from two sources, benign activities (defined as noise), and true threat activities. The *prior estimate of adversary behavior* represented as a series of decisions comprising an attack scenario can be generated in several ways. In our case we use a decision analysis framework that explicitly captures uncertainties on key parameters using decision trees. However, that is not the focus of this paper – for our purposes, the prior estimates can come from

any source, including subject matter experts. The *updating methodology* updates the prior behavior with the cue observations to form a new estimate of adversary behavior. This new estimate can focus resource allocation decisions. The approach is intended to support human analysts in prioritizing steady-state intelligence collection; surveillance or even interdiction assets can be redirected responsively based on new information.

Our primary contribution is a methodology for using the ambiguous cues to update the prior probabilities and generate a posterior probability distribution over possible adversary activity. With a comprehensive scenario structure the evidence from the combination of many corroborating cues, even when ambiguous, can be sufficient to change one's supposition of adversary activity. To accomplish this, we define a class of nuclear attack scenarios and identify possible ambiguous cues with each stage of the scenario.

2. Attack Scenarios

An attack scenario consists of a sequence of actions an adversary would undertake to acquire or fabricate a nuclear weapon and transport it to the US for detonation. In general, an adversary could choose from a large number of potential sources of nuclear weapons or materials, transport paths, fabrication methods, and targets in the US. These choices combine to generate a collection of many possible attack scenarios. We characterize adversary behavior by the relative likelihood of these scenarios.

The major steps that an adversary must successfully execute are summarized in Figure 1. Variations within each of these steps (e.g., source of material, location for weapon fabrication, transportation route) lead to a large number of potential scenarios. A similar outline of necessary steps for a terrorist nuclear attack is described in Bunn et al. (2003). Some additional elements of terrorist attack planning are also outlined in Army (2007). An example illustrating how these high level steps might be broken down into lower level tasks with associated adversary decisions is presented in Table 1. Note that many of the scenario decisions in Table 1 might generate cues prior to the execution of an attack, however we assume that cues for each decision could be observed at the assigned scenario step. Each of the major steps is described as follows.

Adversary leadership, in collaboration with adversary cells that execute the attack, must first plan the attack, recruit operatives, and train them. Execution of the plan starts with procurement of needed materials and equipment. Although much of the material and equipment may be commonplace, some may be unique to nuclear weapons activities. In addition, surveillance of foreign facilities and personnel protecting nuclear weapons or material might be undertaken. Finally, dry runs to probe security or to test the effectiveness of other planned activities may be conducted to prepare for further actions.

Figure 1 **Major steps for an adversary to successfully complete an attack on the US.**

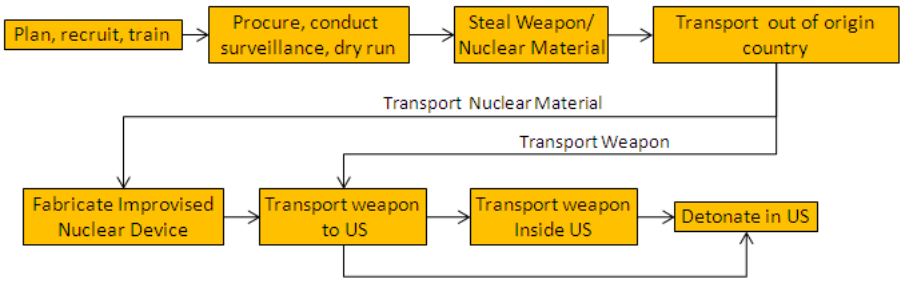


Table 1 **Tasks for each hypothetical scenario step with example scenario decisions**

Scenario Step	Associated Tasks	Example Scenario Decisions
Plan, Recruit, Train	Adversary leadership and cells plan scenario	Determine whether to steal a weapon or material
	Leadership recruits operatives/experts for scenario	Local operatives or foreign operatives
	Operatives train for tasks	
Procure, Conduct Surveillance, Dry Run	Operatives conduct surveillance along source route and target	Entry/Exit ports, route path
	Adversaries conduct procurements for operations	Purchase or acquire transport vehicles
	Adversaries arrange for local and international transportation	Mode of transport: land, sea, air
	Operatives conduct dry run	
Adversary Steals Weapon or Material	Insiders steal material from facility	Fissile material type
	Insider theft of weapon from a facility	Country of origin for weapon
Transport Material/Weapon out of Country	Transport material or weapon from facility to source country exit port	Mode of transport: land, sea, air
		Exit port selection: large/small air/sea port
	Transport material to fabrication country	Country of fabrication
Fabricate Improvised Nuclear Device	Fabricate IND	Type and design of IND to be fabricated
Transport weapon to US	Move IND/Weapon to final exit port for transit to US	Mode of transport: land, sea, air
	Load IND/Weapon at exit port	
	Transport IND/Weapon to US	
Transport weapon within US	Move weapon across US border	Mode of transport: land, sea, air
	Transport weapon to final destination	

In the third step, the adversary attempts to acquire a nuclear weapon or special nuclear material. In our analysis we assume acquisition involves theft of either material or an intact weapon. The theft could occur by the use of force to attack a facility or covertly by insiders who work at a

facility. After the theft, the adversary must transport the weapon or material within the country of origin to the border. Many different transportation routes and modes are possible.

As indicated in Figure 1, if an adversary steals nuclear material, a fabrication step is required. Fabrication of the weapon could take place in the country where the material was stolen, or the material could first be transported to another country. The operational weapon must then be transported through international waters or airspace to the US via one of many possible paths. After penetrating the US border, the weapon is then transported to the target city and detonated. Note that many major cities are at the US border; transportation within the US would be minimal for these targets.

2.1. Cues

In general, cue generation and observation can be characterized by three processes. First, adversary actions may or may not generate an observable cue. Second, if a cue is generated, it may or may not be observed. In our model, we combine these first two factors to reduce the number of likelihood estimate steps required. Finally, observed cues that might be associated with adversary actions may also be due to events that are unrelated to a nuclear threat scenario. We consider cues as noise when generated by benign events. This noise is a primary factor limiting the effectiveness of any assessment.

As an example, consider an adversary using sea transportation to move a weapon. A potential cue for this action is the observation of a suspect terror group purchasing a boat capable of a trans-Atlantic voyage. Acquisition of the boat may or may not generate this purchasing cue – the adversary may be able to use a boat from a sympathizer of the group instead of purchasing a boat. Even if the cue for the boat purchase were generated, it may not be observed. Additionally, there is noise because adversaries may purchase boats for reasons unrelated to a nuclear threat scenario.

The process of identifying plausible cues is aided by the development of hypothetical attack scenarios. Once the major tasks and events for the scenarios are defined, we can determine an initial set of plausible cues. In this initial study, the cues ranged from intercepts of communications on various topics, to secondhand reports, to direct observations. We identified hypothetical cues that could be observed for all of the scenario events. We then consulted with subject matter experts to refine, add to, and validate the initial list.

2.1.1. Impact of cues. Cues are characterized by two probabilities: the probability of detecting the cue given an adversary is performing the tasks in one of the threat scenarios (P_D), and the probability of falsely interpreting the cue when the threat scenario is not underway (P_{FA}). As mentioned previously, the P_D term combines the probability that the cue is generated with the probability the cue is observed when it is generated by threat scenario actions. This represents

Table 2 Impact of Cues as a function of $\frac{P_D}{P_{FA}}$ and P_D

Cue Type	Likelihood of Seeing Cue (P_D)	$\frac{P_D}{P_{FA}}$	Impact of cue seen	Impact of cue not seen	Example
1	High	Strong	Very high	Very high	Direct intercept of adversary comm. referencing nuclear-specific terms, expertise, locations, transport logistics.
2	High	Weak	Low	High	Attempts to recruit nuclear expertise.
3	Low	Strong	Very high	Low	Source country notifies US of missing weapon.
4	Low	Weak	Low	Very low	Member of suspect group purchases small boat.

the overall likelihood of observing a valid cue associated with an adversary task. The P_{FA} can be thought of as a noise term associated with cues that are generated by activities unrelated to threat scenarios. It is useful to consider the ratio $\frac{P_D}{P_{FA}}$ along with P_{FA} because is a measure of the signal to noise ratio, giving a sense of how strong an observed cue can be in support of a hypothesized threat scenario. The value for the noise term P_{FA} is easily calculated from these two values: $P_{FA} = \frac{P_D}{P_D/P_{FA}}$.

In Table 2 we show how the two values, P_D and P_{FA}, interact and the impact they can have on the belief of adversary behavior. Four different types of cues are described in the table. The first type corresponds to information that is highly specific and closely monitored, so it provides useful evidence whether it is observed or not. We assume that the failure to observe a cue in a data stream that is closely monitored provides evidence that the threat scenario is not underway. The second type of cue is not very definitive, so its contribution to the corpus of evidence is low when it is observed. However, not seeing the cue is a significant indicator that the activity is not occurring. The third type of cue is rarely observed but highly specific. Impacts are high if it is observed and low when it is not observed. Finally, the fourth type is neither closely monitored nor definitive. Such cues have small impacts. The mathematical formalism for analysis using these cues is developed in Section 3.

2.1.2. Scoring of cues. One of the most difficult tasks in developing the cues for analysis is scoring the cues – i.e., determining the P_D and P_{FA} of each cue. We assume that the values of P_D for each cue could be provided by subject matter experts. However, the values for P_{FA} could be approximated by examining historical data streams and observing how often a given cue appears. Assuming no nuclear threat scenarios were underway in the time frame of the data, these cues would correspond to false alarms.

If subject matter experts were reluctant to provide quantitative values for these probabilities, they could be asked to assess the strength of cues qualitatively for this study. That is, cues could be characterized as "strong," "moderate," or "weak." Other terms such as "definitive," "near definitive," or "less definitive," or other categories, could also be used. For this study, qualitative terms were mapped into numerical values of P_D or $\frac{P_D}{P_{FA}}$ for analysis of this hypothetical example.

Since these assessments are complex and subjective, there will likely be significant uncertainty related to the values of P_D. In Section 5, we describe an initial sensitivity analysis that was done to assess the robustness of the model to these uncertainties.

2.1.3. Example set of cues associated with an attack scenario. We provide an example list of cues for a hypothetical attack scenario in Table 3. The cues are associated with tasks and decisions that may be made as suggested by Table 1. These cues may not necessarily be observed at the same time as the actual task. In this table, the strength of each cue is specified in terms of qualitative values for $\frac{P_D}{P_{FA}}$ and P_D.

3. Updating scenario event probabilities given cues

The starting point for the modeling framework to exploit ambiguous cues is a prior probabilistic model of adversary behavior. This prior model of adversary behavior can be elicited from experts or be the output of a separate model characterizing adversary behavior. One such approach uses a decision tree analysis (Maurer (2009)) that is extended with Monte Carlo simulation of uncertain parameters. In this formulation, the optimal adversary decision is a probability distribution over possible actions rather than a deterministic choice at each node. By doing so we are able to use previously developed prior models to feed into the event tree analyzed in this section. The focus of this paper is the cue updating mechanism for revising these prior probabilities.

We use an approach based on event trees and Bayes' Rule to incorporate the evidence from cues into the decision tree framework. We considered related modeling approaches, including hidden Markov models (Rabiner and Juang (1986), Rabiner (1989)) and general Bayesian networks (Heckerman (1995)). However, neither of these approaches exploits the scenario event tree structure present in our problem.

A hidden Markov model (HMM) was considered since the true state of the adversary is unknown and cue observations are generated by one or more hidden states. In a hidden Markov model, an observer of the system can only see the output of hidden states and thus cannot unambiguously associate observables with a specific hidden state. At no time does the observer know the current state of the system. In this case the system states represent the adversary activities and the cues are the observations. One common approach to solving the HMM problem of determining the original activities given observations is with the Viterbi algorithmm. However, this will only yield

Table 3 Example Set of Hypothetical Scenario Cues with Associated Tasks and Strengths

Cue	Associated Task	Nominal $\frac{P_D}{P_{FA}}$	Nominal P_D
Comm[1]: Plan to steal material	Plan material theft	Weak	Low
Elevated comm traffic at material storage facility	Steal material	Weak	Low
Comm: Material has been succesfully procured	Steal material	Moderate	Very Low
Comm: Plan to steal weapon	Plan weapon theft	Weak	Low
Security mobilized in source state weapons complex	Steal wapon	Weak	Low
Source state government reports stolen weapon	Steal weapon	Very Strong	Very Low
Boat purchase by suspect	All sea transit events	Weak	Low
Comm: Boat use for sensitive package transportation	All sea transit events	Weak	Low
Local report of suspicious foreigner loading canister at source country seaport	Sea Transport to Fabrication	Weak	Low
Local report of suspicious foreigner loading canister at source country airport	Air Transport to Fabrication	Weak	Low
Elevated safe house activity in Country A	Fabricate in Country A	Weak	Low
Comm: Source state to Country A: equipment/expert required	Fabricate in Country A	Weak	Low
Secondhand allegation of fabrication in Country A	Fabricate in Country A	Weak	Low
Elevated Country B safe house activity	Fabricate in Country B	Weak	Low
Comm: Source state to Country B: equipment/expert required	Fabricate in Country B	Weak	Low
Secondhand allegation of fabrication in Country B	Fabricate in Country B	Weak	Low
Comm: US seaport to be used	Sea Transport to US	Weak	Very Low
Private plane purchase by suspect	All Air Events	Weak	Very Low
Suspicious documentation on international flight to US	Air Transport to US	Weak	Low
Unidentified aircraft approaching US	Air Transport to US	Weak	Low
Comm: POE[2] Entry City targeted	POE Entry	Weak	Low
Comm: Non-POE entry to be used	Non-POE entry	Weak	Low
US locals reports non-POE entry by suspicious persons	Non-POE entry	Weak	Low
Unexplained radiation detected at domestic road monitors	Road to Detonation	Weak	Low
Unexplained radiation detected in train cargo	Rail to Detonation	Weak	Low

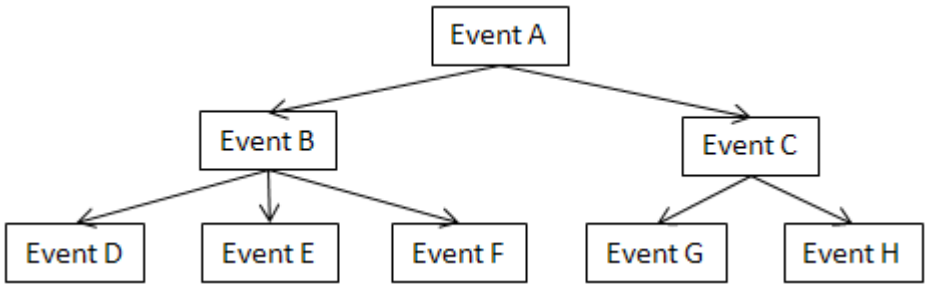
[1]Comm: Intercepted communication transmissions indicating given action. [2]POE: Port of entry

the maximum likelihood path (Rabiner and Juang (1986), Rabiner (1989)). In our analysis, we seek the probability distribution over all possible paths, not just the most likely; this is more useful in informing resource allocation decisions.

Additionally, since in our problem we do not have repeated states or cycles that a hidden Markov model can model, a tree structure is more appropriate for representing adversary states. Additionally, the tree structure enables a much more computationally efficient way of calculating posterior probabilities.

Our approach closely resembles a Bayesian network where inference is performed on directed acyclic graphs. However, unlike Bayesian networks in general, the problem considered in this study has a special structure that we are able to exploit. In our scenario event tree, the probabilities

Figure 2 Example event tree.



of all branches emanating from each event node sum to one; this is not true for general Bayesian networks. Consequently, we are able to employ a specialized solution methodology. Section 3.1 describes the tree structure of our approach as well as the methodology used to capture the evidence provided by the cues.

3.1. Statistical updating mechanism based on priors

The goal of the updating methodology employed is to answer the question: Given a set of cues, some of which have been observed, what is the probability that the adversary is following a particular course of action as described by the scenario tree?

Our updating methodology recursively applies Bayes' Rule to incorporate the prior path probabilities along with the cue likelihoods that were assessed with assistance from subject matter experts. In our study, we assume that a single attack scenario is occurring, and the goal of this method is to convey the relative probabilities of occurrence of different scenarios. To illustrate the updating mechanism, consider the tree in Figure 2.

Each branch split has an associated prior probability; this prior probability does not account for any information given by the cues. For example, event A may represent an adversary deciding to steal nuclear material. Event C may represent the choice of transporting the material to a second country via sea. If the decision analysis (or subject matter expert) indicates that the adversary has a 60% chance of transporting material by sea after stealing nuclear material, then $P(C|A) = 0.6$.

Each event in the tree may have one or more associated cues. A cue for the example above may be that a member of a terrorist group purchased a boat capable of a transatlantic voyage.

We define the notation used in the remainder of this section.

- $P(A)$: denotes the probability of adversary action for event A
- $P(\bar{A})$: denotes the probability of NO adversary action for event A
- $P(AC)$: denotes the probability of adversary action for event A followed by adversary action for event C
- $P(\overline{AC})$: denotes the probability of adversary action in NEITHER event A nor event C

- a_i: denotes the observation of the ith cue corresponding to event A. If there is only one cue corresponding to event A, the subscript is dropped.
- $P(C|A)$: denotes the probability that event C occurs given that event A has occurred. This expression does not presume that the defender is aware that A has occurred. This is the result from the prior model of adversary behavior that is either generated by a a separate decision analysis model or elicited from experts.
- $P(a_i|A)$: denotes the probability that cue a_i is generated and observed given adversary action for event A
- $P(a_i|\bar{A})$: denotes the probability that cue a_i is generated and observed given NO adversary action for event A (i.e., a false positive)
- $P(cues|A,C)$: shorthand for the probability of observing a set of cues given adversary action for events A and C.

The following assumptions are also made:

Assumption 1. $P(G|A,C) = P(G|C)$: The probability that event G occurs given that event C occurred does not depend on the fact that event A also occurred prior to event C.

Assumption 2. $P(a_i, b_i|A,B) = P(a_i|A,B)P(b_i|A,B) = P(a_i|A)P(b_i|B)$: Cues are not correlated with each other, and the probability of observing a cue depends only on the corresponding adversary event and not on the path taken to reach that event.

We begin by describing the updating mechanism where each event has only one possible cue that could be observed (i.e. event A has cue a, event B has cue b, etc.). Referring to Figure 2, we ask the question: given a set of cues which we either observe or do not observe, what is the probability the adversary is in, or has progressed beyond event G? It is sufficient to ask what is $P(G|cues)$. By the law of total probability then:

$$P(G|cues) = P(G|A,C,cues)P(A,C|cues) + P(G|\overline{A,C},cues)P(\overline{A,C}|cues)$$

From the tree, we immediately see that we can ignore the second term in the equation (it is equal to zero) since event G cannot occur without events A,C first occurring. This means that the equation simplifies to:

$$P(G|cues) = P(G|A,C,cues)P(A,C|cues) \tag{1}$$

Applying Bayes' Rule to $P(G|A,C,cues)$ gives:

$$P(G|A,C,cues) = \frac{P(cues|A,C,G)P(G|A,C)}{P(cues|A,C)}$$

which expands to

$$P(G|A,C,cues) = \frac{P(cues|A,C,G)P(G|A,C)}{P(cues|A,C,G)P(G|A,C) + P(cues|A,C,\bar{G})P(\bar{G}|A,C)}$$

by the law of total probability. By assumption 1 and the structure of the tree, it is equivalent to

$$P(G|A,C,cues) = \frac{P(cues|A,C,G)P(G|C)}{P(cues|A,C,G)P(G|C) + P(cues|A,C,H)P(H|C)} \tag{2}$$

The term $P(G|C)$ in equation 2 is the result from the prior model of adversary behavior that is either generated by a separate decision analysis model or elicited from experts.

The cue set likelihood $P(cues|A,C,G)$ in equation 2 is the probability of observing the cues given that the adversary is taking path A,C,G. For example, if we observe the cues a, b, and g but none of the other cues, the cue set is $\{a,b,\bar{c},\bar{d},\bar{e},\bar{f},g,\bar{h}\}$. By Assumption 2 we can re-write this expression as

$$P(\{a,b,\bar{c},\bar{d},\bar{e},\bar{f},g,\bar{h}\}|A,C,G) = P(a|A)P(b|\bar{B})P(\bar{c}|C)P(\bar{d}|\bar{D})P(\bar{e}|\bar{E})P(\bar{f}|\bar{F})P(g|G)P(\bar{h}|\bar{H}) \tag{3}$$

Note that the terms on the right side of equation 3 are precisely those that we need to assess from subject matter experts.

We return to equation 1 which can now be re-written as:

$$P(G|cues) = \frac{P(cues|A,C,G)P(G|C)}{P(cues|A,C,G)P(G|C) + P(cues|A,C,H)P(H|C)} P(A,C|cues) \tag{4}$$

by substituting equation 2 into equation 1, where the only unknown remaining is $P(A,C|cues)$. To solve for $P(A,C|cues)$ we apply the same procedure as we did for $P(G|cues)$. That is:

$$\begin{aligned} P(A,C|cues) &= P(C|cues) \\ &= P(C|A,cues)P(A|cues) \cancel{+P(C|\bar{A},cues)P(\bar{A}|cues)} \\ &= P(C|A,cues)P(A|cues) \end{aligned}$$

where the second term in the equation is zero due to the structure of the tree. We proceed as we did in equation 2.

$$P(C|cues,A) = \frac{P(cues|A,C)P(C|A)}{P(cues|A)}$$

where:

$$\begin{aligned} P(cues|A) &= P(cues|A,C)P(C|A) + P(cues|A,\bar{C})P(\bar{C}|A) \\ &= P(cues|A,C)P(C|A) + P(cues|A,B)P(B|A) \end{aligned}$$

Note that $P(cues|A,C)$ has already been computed as the denominator in equation 2. Finally we have:

$$P(A,C|cues) = \frac{P(cues|A,C)P(C|A)}{P(cues|A,C)P(C|A) + P(cues|A,B)P(B|A)} P(A|cues)$$

Note that all of the terms in the above equation are now known. Since A is the root node, $P(A|cues) = 1$. In this way, the posterior probabilities for the entire tree are calculated by recursively applying the above methodology throughout the tree. Each computed posterior propagates down the tree to affect the computation of posteriors further down the tree.

3.1.1. More complex tree structures. The method described above presents the computations for a basic binary tree structure and cue observations. We also consider all of the different more complex tree structures.

More than two branches from a node. If there are more than two branches from a particular node, then the denominator of equation 2 incorporates all of the branches from that event. For example, there are three branches emanating from event B in Figure 2. Thus:

$$P(cues|A,B) = P(cues|A,B,D)P(D|A,B) + P(cues|A,B,\bar{D})P(\bar{D}|A,B)$$

Which again, as dictated by the tree structure, simplifies to:

$$P(cues|A,B) = P(cues|A,B,D)P(D|B) + P(cues|A,B,E)P(E|B) + P(cues|A,B,F)P(F|B)$$

The rest of the procedure remains unchanged.

Multiple cues per event. If there are multiple cues associated with one event, the procedure remains much the same. The calculation of the cue likelihood $P(cues|Path)$ will incorporate this possibility. For example, if event C has two cues, $\{c_1, c_2\}$, then there will also be $P(c_1|C)$ and $P(c_2|C)$. Equation 3 becomes:

$$\begin{aligned} P(\{a,b,\bar{c_1},\bar{c_2},\bar{d},\bar{e},\bar{f},g,\bar{h}\}|A,C,G) = {} & P(a|A)P(b|\bar{B})P(\bar{c_1}|C)P(\bar{c_2}|C)P(\bar{d}|\bar{D}) \\ & P(\bar{e}|\bar{E})P(\bar{f}|\bar{F})P(g|G)P(\bar{h}|\bar{H}) \end{aligned}$$

This probability can then be used in equation 4 with the rest of the analysis unchanged.

Same event appears at different locations in the tree. There may be many cases in a larger tree where two different nodes on the tree represent the same event with the same cues where the only difference is the path taken to reach that event. For example, event D and event G could represent the adversary fabricating a weapon in Country A. Events D and G represent the same event and can generate one specific cue, represented as cue d. Therefore we can call event G event D. For example equation 3 becomes:

$$P(\{a,b,\bar{c},d,\bar{e},\bar{f},\bar{h}\}|A,C,D) = P(a|A,C,D)P(b|A,C,D)P(\bar{c}|A,C,D)P(d|A,C,D)P(\bar{e}|A,C,D)P(\bar{f}|A,C,D)p(\bar{h}$$
$$P(\{a,b,\bar{c},d,\bar{e},\bar{f},\bar{h}\}|A,C,D) = P(a|A)P(b|\bar{B})P(\bar{c}|C)P(d|D)P(\bar{e}|\bar{E})P(\bar{f}|\bar{F})p(\bar{h}|\bar{H})$$

Similarly if the equivalent states are not in consideration by the $Path$:

$$\begin{aligned} P(\{a,b,\bar{c},d,\bar{e},\bar{f},\bar{h}\}|A,C,H) &= P(a|A)P(b|\bar{B})P(\bar{c}|C)P(d|\bar{D})P(\bar{e}|\bar{E})P(\bar{f}|\bar{F})p(\bar{h}|H) \\ &= P(a|A)P(b|\bar{B})P(\bar{c}|C)P(\bar{e}|\bar{E})P(\bar{f}|\bar{F})p(d|\bar{G})p(\bar{h}|H) \end{aligned}$$

The remaining analysis follows as before.

3.1.2. Role of cue strength and likelihood in updating mechanism. There are several features of this approach that are noteworthy. As previously stated, the ratio of the probability of observing a cue given a threat event, P_D, to the probability of observing a cue given a benign event, P_{FA}, is one measure of the strength of a cue. The cause for this is rooted in the Bayes' equation. Consider the form of equation 2 with the expanded marginal normalization factor:

$$P(G|A,C,cues) = \frac{P(cues|A,C,G)P(G|C)}{P(cues|A,C,G)P(G|C) + P(cues|A,C,\bar{G})P(\bar{G}|C)} \tag{5}$$

This can be rewritten as:

$$P(G|A,C,cues) = \frac{P(G|C)}{P(G|C) + \frac{P(cues|A,C,\bar{G})}{P(cues|A,C,G)}P(\bar{G}|C)}$$

In this case, if we consider the cue g for state G, then the fractional term in the denominator can be interpreted as:

$$\frac{P(g|\bar{G})}{P(g|G)} = \frac{P_{FA}}{P_D}$$

This is the reciprocal of the ratio of the probability of observing a cue given a threat to the probability of observing the cue given no threat. $\frac{P_D}{P_{FA}}$ is also one way to express the Bayes' factor (Kass and Raftery (1995)). So if the cue is "stronger," then the ratio $\frac{P(g|\bar{G})}{P(g|G)} = \frac{P_{FA}}{P_D}$ is smaller, which in turn increases the value of $P(G|A,C,cues)$.

Additionally, the lack of observing a cue influences the probability that the adversary takes a specific path. To determine the impact of not observing a cue that you are looking for, consider a single unobserved cue for state $G, \bar{g}$. We can then write equation 5 as:

$$P(G|A,C,\bar{g}) = \frac{P(\bar{g}|A,C,G)P(G|C)}{P(\bar{g}|A,C,G)P(G|C) + P(\bar{g}|A,C,\bar{G})P(\bar{G}|C)}$$

But since $P(\bar{g}|G) = 1 - P(g|G)$ and $P(\bar{g}|\bar{G}) = 1 - P(g|\bar{G})$ we get:

$$P(G|A,C,\bar{g}) = \frac{P(G|C)}{P(G|C) + \frac{1-P(g|A,C,\bar{G})}{1-P(g|A,C,G)}P(\bar{G}|C)}$$

where the important ratio to consider is:

$$\frac{1-P(g|A,C,\bar{G})}{1-P(g|A,C,G)} = \frac{1-P_{FA}}{1-P_D}$$

Here, it is easier to interpret the effect of not observing a cue in terms of its P_D, which is $P(g|G)$. The higher the P_D of a cue, the greater the impact will be of not observing that cue. To illustrate, consider the example if $P(g|G) = 0.9$ and $P(g|\bar{G}) = 0.2$, then $\frac{1-P_{FA}}{1-P_D} = 8$. This means that not seeing this cue will decrease the path probability significantly. However, if $P(g|G) = 0.09$ and $P(g|\bar{G}) = 0.02$, then $\frac{1-P_{FA}}{1-P_D} = 1.08$. In this case, *not seeing* the cue will only slightly decrease the path probability. By contrast, the impact of *seeing* the cue on the path probability is the same since the Bayes' factor $\frac{P_D}{P_{FA}}$ is the same in both cases.

Since both observing and not observing a cue affects the posterior probabilities, the observation of a cue can affect the posterior probabilities for all scenario paths.

4. Effect of Cues on Adversary Probability Updating

In this section we show how ambiguous cues can inform an analyst of a possible adversary attack and that the formalism is consistent with the intuition of the analyst. To accomplish this, we first define what kind of properties a cue exploitation model must have. We then introduce an example event tree that uses the cues in Table 3 to illustrate exactly these properties. We assign default cue strength values to these cues, and we select specific locations on the tree to observe changes in posterior probabilities. Then we present several examples of observing cues in various configurations that highlight how this formalism is consistent with the intuition of the analyst.

For a cue exploitation method to be useful, it must have the following effects on the prior belief about adversary behaviors:

1. Observing a weaker cue (i.e. one with a lower $\frac{P_D}{P_{FA}}$) will have less of an effect on the belief about adversary decisions than observing a stronger cue (i.e. one with a higher $\frac{P_D}{P_{FA}}$).

2. Not observing a cue can affect belief about adversary decisions, especially when the non-observed cue has a high P_D.

3. Observing many weak cues scattered across the event tree has little effect on the belief about adversary decisions.

4. Observing a few weak cues along one scenario path can have some impact on the belief about adversary decisions.

5. Observing many weak cues along one scenario path can have a significant impact on the belief about adversary decisions.

6. Observing a cue for an event should impact belief about upstream and downstream events

Our example tree incorporates all of the major steps for an adversary attack on the US as described in Figure 1. The full event tree for the following examples is compacted and split into a top and a bottom subtree as shown in Figures 3 and 4. The bottom subtree is repeated for each "Air/Sea Transit to US" event at the bottom of the tree in Figure 3. As previously discussed, the priors on the underlying event tree could be given by the decision analysis model or by subject matter experts. In the example used in this section, the priors are assumed given to illustrate key points of the updating approach.

There are two main adversary strategies that are modeled in this example: the steal material and steal weapon scenario paths. The main difference between the two is that the steal material scenario paths must include a nuclear device fabrication step. The product of this fabrication step is an Improvised Nuclear Device (IND).

To show the full tree represented by Figures 3 and 4 in one figure, we collapse several *Transit to US* event sub-trees. All of the *Transit to US* sub-trees share the same cues, so if there is a *Transit to US* cue shown in the following example trees, then it also appears in the collapsed event

Figure 3 **Detailed view of the top of the example event tree. The events following each "Air/Sea Transit to US" event are shown in Figure 4 as indicated by the dotted lines at the bottom of the tree. The dotted arrows indicate that there are other possible events following each node that are not modeled in our example.**

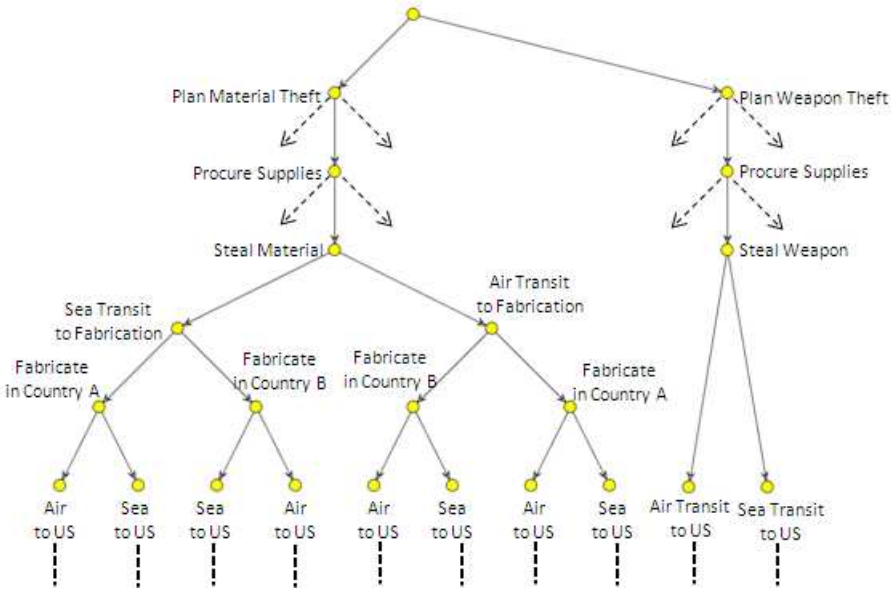
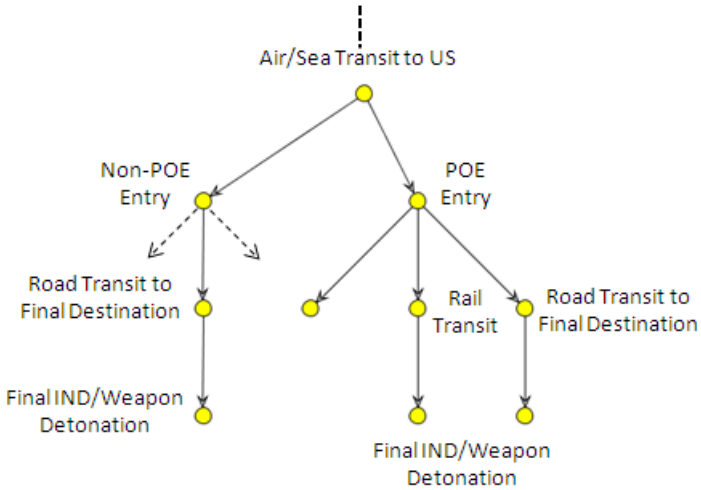


Figure 4 **Detailed view of the event sub-tree following each "Air/Sea Transit to US" event. The Non-POE entry branch was reduced to one path for simplicity, although it could be expanded with more adversary events. The final stage in each path is referred to as "Final IND/Weapon Detonation." In cases where material is stolen, the final event node label is actually "Final IND Detonation." In cases where a weapon is stolen, the final event node label is "Final Weapon Detonation."**



sub-trees that are not shown. Figure 5 shows a reduced form of the tree. We will refer to the events in this figure to identify where cues are observed in the examples.

Figure 5 **A collapsed version of the tree represented by Figures 3 and 4. The dotted lines indicate subtrees that have been omitted in order to visualize the tree on paper. Highlighted in circles are six locations within the event tree which we will focus our attention to changes in posterior probabilities relative to the prior probabilities.**

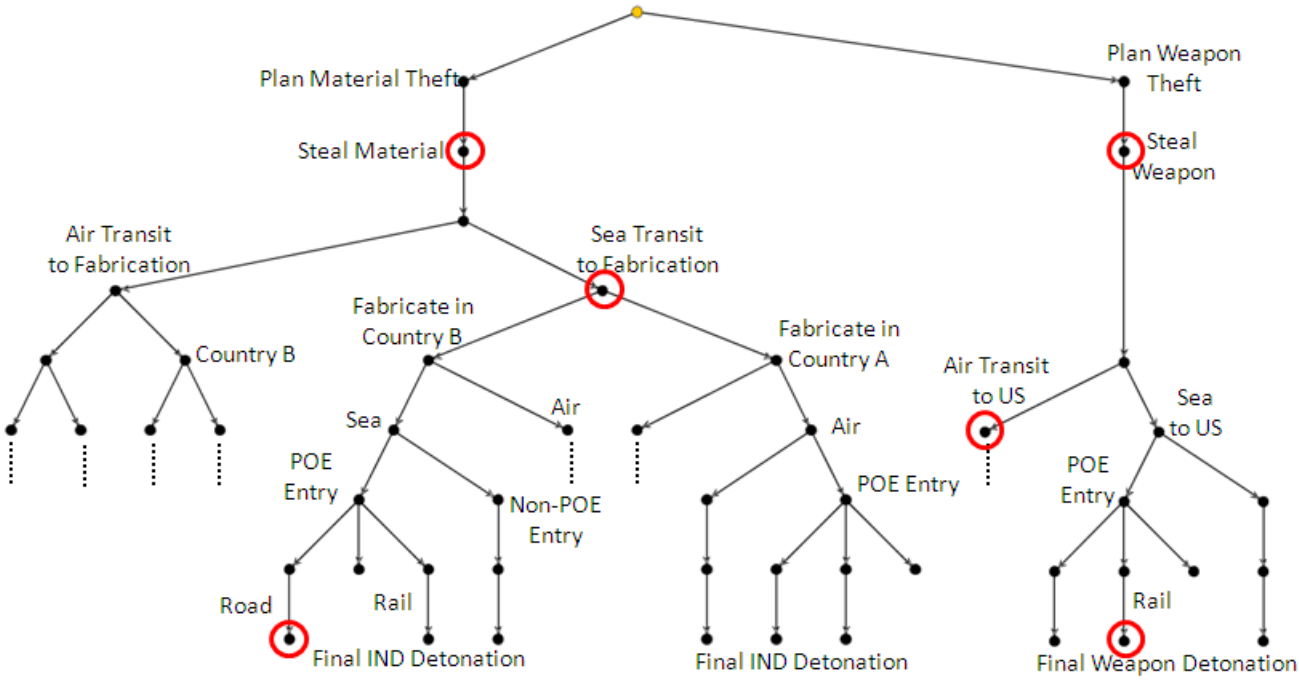


Table 4 **Default cue strengths used in the example tree for the cues listed in Table 3**

Cue	Default P_D	Default P_{FA}	$\frac{P_D}{P_{FA}}$
Comm: Material has been succesfully procured	0.05	0.025	2
Source state government reports stolen weapon	0.01	0.001	10
Unexplained radiation detected at domestic road monitors	0.05	0.04	1.25
Unexplained radiation detected in train cargo	0.05	0.04	1.25
All other cues in Table 3	0.1	0.08	1.25

The cues listed in Table 3 and their associated tasks correspond to the events in this example event tree. Note that there are cues that will appear in multiple locations across the tree if observed. For demonstrative purposes, we assign default values of P_D and P_{FA} to correspond to the notional cue strengths listed in the table. These values are not reflective of actual assessed values. For the majority of events which have a "Low" P_D and a "Low" $\frac{P_D}{P_{FA}}$, we assign $P_D = 0.1$ and $P_{FA} = 0.08$ giving $\frac{P_D}{P_{FA}} = 1.25$. For other cues that do not have these strength attributes, we assign default values of P_D and P_{FA} as indicated in Table 4.

Table 5 Prior probabilities of selected events in the event tree.

Event	Path Prior Probability
Steal Material	0.40
Steal Weapon	0.60
Sea Transit Material to Fabrication Country	0.24
Air Transit Weapon To US	0.36
Final IND Detonation	0.01
Final Weapon Detonation	0.04

We pick six locations within the event tree to observe the changes in posterior probabilities relative to the prior probabilities. These nodes are highlighted with red circles in Figure 5 and correspond to the following actions:

For the material path:

- Adversary plans to steal material and successfully does so.
- Adversary steals material and uses sea transit to the fabrication country.
- Adversary steals material, uses sea transit to fabricate in Country B. Then the adversary uses sea transit to a large US border city, and uses road transportation to the final target and detonates the IND.

For the weapon path:

- Adversary plans to steal a weapon and successfully does so.
- Adversary steals a weapon and uses air transit to enter the US.
- Adversary steals a weapon and uses sea transit to a large US border city, and then uses rail transportation to the final target and detonates the weapon

Table 5 shows the prior probabilities for the selected events in the event tree. These represent the probability an adversary will choose to follow the path that reaches the corresponding event. The prior probabilities can be derived from the decision tree model of adversary behavior or based on expert opinion, but are only notional for this example.

The following examples are meant to capture qualitative behaviors of the modeling and updating mechanism. Since there are many confounding variables and parameters when considering cues and their effect on belief, we illustrate properties of the methodology for a few representative examples.

4.1. Effect of observing a single cue depends on cue's strength

Observing a single cue will have a varying effect on our belief about adversary decisions depending upon its strength. That is, an observed weak cue will have little effect on the prior belief, while an observed strong cue can have a significant effect. The methodology does follow this intuitive reasoning.

Table 6 The posterior probability of an event as a function of the cue strength that is observed at the event given a biased prior against that event.

$\frac{P_D}{P_{FA}}$	$p(Steal\ Weapon \mid cues)$
1	0.001
10	0.01
100	0.10
870	0.50
1000	0.54

To illustrate the effect of cue strength, consider an example where there is one branch with an extremely low prior. This also illustrates how strong an observed cue must be in order to overcome this strong bias in the tree. We focus on the *Steal Weapon/Steal Material* branches of the event tree in Figure 5. We set the prior probability of *Steal Weapon* to be low at $p = 0.001$. We observe one cue signifying a stolen weapon, the "Source state government reports stolen weapon" cue and vary its strength as defined by $\frac{P_D}{P_{FA}}$.

We vary the cue strength ratio in order to determine the point at which the *Steal Weapon* scenario event becomes more likely than the *Steal Material*. Table 6 shows the probability of the steal weapon branch for different values of $\frac{P_D}{P_{FA}}$. Note that a strength of $\frac{P_D}{P_{FA}} = 1$ indicates that the cue will have no effect on the prior probability.

The probability for the steal weapon branch with an extremely low prior probability requires a very strong cue to switch our belief in the adversary activity away from the material branch. With a prior probably of 0.001, a cue strength close to 1000 (870 is the approximate threshold) is required to favor this path. This stems from the fact that equation 5 can be reformulated to show that the posterior is dependent on the Bayes' factor multiplied by the event prior.

4.2. Not observing a cue can affect path probabilities

Not observing a cue can have an effect on the posterior path probabilities. This is important in cases where a cue is expected with high probability to be observed if a certain event occurs, i.e., the cue has a high P_D. To demonstrate this, consider the cue of "Elevated safe house activity in Country B." In our example, suppose that if a weapon is being fabricated in Country B, there would be a very high chance that this cue would be observed. To model this we set the $P_D = 0.7$. We keep the Bayes' factor strength at the default of 1.25 by setting $P_{FA} = 0.56$.

Table 7 shows that any paths that include *Fabrication in Country B* have a lower posterior probability. As expected, this decreases our belief that the material theft scenario is underway. As a result, the belief that the weapon theft scenario is underway increases since we assume that at least one of the scenarios is underway.

Table 7 Difference between path posteriors and priors when not observing a cue with high P_D.

Event	Path Prior	Path Posterior	Difference
Steal Material	0.40	0.33	-0.07
Steal Weapon	0.60	0.67	0.07
Sea Transit Material to Fabrication Country	0.24	0.20	-0.04
Air Transit Weapon To US	0.36	0.40	0.04
Final IND Detonation	0.01	0.01	0.00
Final Weapon Detonation	0.04	0.04	0.00

4.3. Many observed weak cues scattered across the event tree have little effect

If we examine a case of observing many weak cues scattered across many branches, we would not expect much change compared to prior path probabilities since the observation of a cue in one location will likely be offset by the observation of a cue in another location on the event tree. To illustrate this, we select several weak cues associated with events scattered throughout the tree. For this example, nine different weak cues with $P_D = 0.1$ and $P_{FA} = 0.08$ were observed while the remaining cues were not observed. The cues observed in this example are listed below next to their associated events:

- *Steal Material*: Elevated comm traffic at Material storage facility
- *Steal Weapon*: Security mobilized in source state weapons complex
- *Fabricate in Country A*: Secondhand allegation of fabrication in Country A
- *Fabricate in Country B*: Elevated Country B safehouse activity
- *POE Entry*: Comm: POE Entry City targeted
- *Non-POE entry*: Comm: Non-POE entry to be used, US locals reports non-POE entry by suspicious persons
- *Road to Detonation*: Unexplained radiation detected at domestic road monitors
- *Rail to Detonation*: Unexplained radiation detected in train cargo

These cues are scattered in a relatively uniform manner across the event tree. The posterior path probabilities for several event paths are given in Table 8. There was very little change in the belief about adversary activities for the most of the event paths.

4.4. Observation of a few weak cues on one scenario path has some impact

One potential value of this methodology is the ability to exploit multiple cues (even when ambiguous) that are observed along the same scenario path. The observation of such aligned cues should increase the belief that the adversary is taking that path. To illustrate, we first present a case where there are a small number of cues along one path in the scenario tree. These cues also correspond to other events across the event tree. The cues selected all have $P_D = 0.1$ and $P_{FA} = 0.08$. The observed cues and their associated states in this example are:

Table 8 Difference between path posteriors and priors when observing many weak cues scattered across the tree.

Event	Path Prior	Path Posterior	Difference
Steal Material	0.40	0.43	0.03
Steal Weapon	0.60	0.57	-0.03
Sea Transit Material to Fabrication Country	0.24	0.25	0.01
Air Transit Weapon To US	0.36	0.34	-0.02
Final IND Detonation	0.01	0.01	0.00
Final Weapon Detonation	0.04	0.03	-0.01

Table 9 Difference between path posteriors and path priors when observing few weak cues in one scenario path.

Event	Path Prior	Path Posterior	Difference
Steal Material	0.40	0.48	0.08
Steal Weapon	0.60	0.52	-0.08
Sea Transit Material to Fabrication Country	0.24	0.32	0.08
Air Transit Weapon To US	0.36	0.28	-0.08
Final IND Detonation	0.01	0.01	0.00
Final Weapon Detonation	0.04	0.04	0.00

- *All sea transit events*: Boat purchase by suspect
- *Fabricate in Country B*: Elevated Country B safehouse activity, Comm: Source state to Country B: equipment/expert required

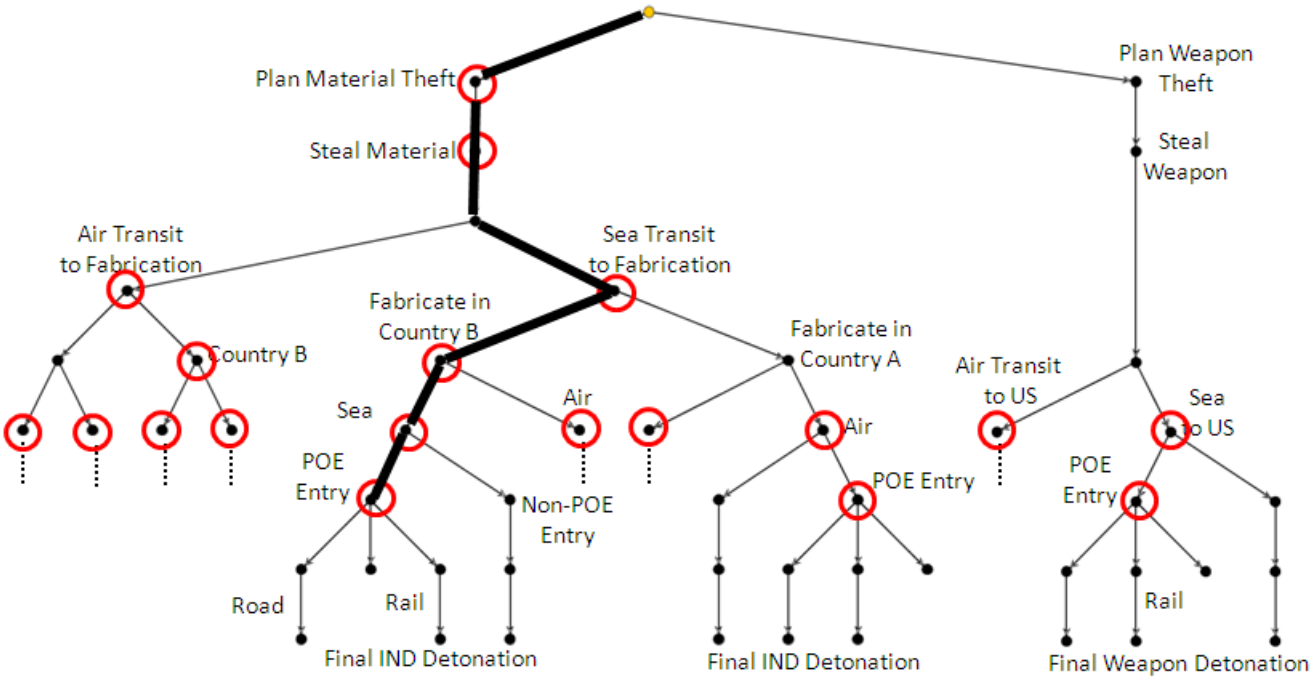
Note that these cues also appear in multiple places in the tree. Table 9 shows the posterior path probabilities compared to the prior path probabilities. The maximum difference in the posterior probability is 0.08 in the *Steal Material* versus *Steal Weapon* events of the tree. This illustrates there is a small but noteworthy change in posterior probability when just a few weak cues are observed along a scenario path.

4.5. Observing many weak cues along one scenario path has a significant impact

If we increase the number of weak cues that are observed on the same scenario path, there is a larger impact on the posterior probabilities. The scenario path the cues highlight is that an adversary steals material transports it to country B using sea transit for a weapon fabrication and uses sea transit to transport the weapon to a port of entry city. Figure 6 highlights this scenario in the example tree. The cues that are involved are:

- *Plan Material Attack*: Comm: Plan to steal material
- *Steal Material*: Elevated comm traffic at material storage facility
- *All sea transit events*: Boat purchase by suspect
- *Sea Transport to Fabrication*: Local report of suspicious foreigner loading canister at source country seaport

Figure 6 **The example collapsed tree with many weak cues observed on the same scenario path. A circle around an event indicates that one or more cues have been observed for that event. The thick line represents the highlighted scenario path. The cues in the highlighted scenario path also are observed elsewhere in the tree.**



- *Fabricate in Country B*: Elevated Country B safe house activity, Comm: Source state to Country B: equipment/expert required, Secondhand allegation of fabrication in Country B
- *Air Transport to US*: Suspicious documentation on international flight to US
- *POE Entry*: Comm: POE Entry City targeted

Note that some of these cues do also appear in other places in the event tree, and there are also many other events not on the scenario path for which cues are observed. Table 10 shows sample posterior probabilities compared to their prior probabilities for the case when there are many cues observed on the same path.

In this case, the *Steal Material* event path has a change of 0.27 over the prior compared to 0.08 in the case with just a few weak cues. This represents a significant increase in probability over the prior compared to the small increase in the previous case. Thus, the more cues observed along the same path the higher our belief that the adversary chose that path, even if cues are observed for events not on the scenario path.

The above example illustrates the need for combining cues from disparate sources. Ambiguous cues that are associated with events at the top level of the tree (e.g., related to safeguards and security) may not ordinarily be available to organizations observing cues being generated at the bottom of the tree (e.g., local law enforcement). Similarly entities collecting ambiguous cues for

Table 10 Difference between path posteriors and path priors when observing many weak cues in a scenario path.

Event	Path Prior	Path Posterior	Difference
Steal Material	0.40	0.67	0.27
Steal Weapon	0.60	0.33	-0.27
Sea Transit Material to Fabrication Country	0.24	0.47	0.23
Air Transit Weapon To US	0.36	0.20	-0.16
Final IND Detonation	0.01	0.03	0.02
Final Weapon Detonation	0.04	0.02	-0.02

events at the bottom of the tree, may be unaware of any corroborating cues associated with top and middle level events. In these cases, a limited set of ambiguous cues may be discarded or ignored. Combining cues from disparate sources that span a larger portion of the scenario paths can provide much more compelling evidence.

5. Sensitivity Analysis

As discussed in Section 2.1.2, it may be difficult for subject matter experts to estimate probabilities of observing cues. Thus, the assessed values of the cue detection probabilities, P_D and P_{FA}, can have significant uncertainty. In order to study the impact of this uncertainty on the posterior probabilities of cue uncertainty we perform an empirical sensitivity analysis. The purpose of this analysis is to determine the impact on the posterior when selecting the P_D or $\frac{P_D}{P_{FA}}$ values for different qualitative categories of cue strength.

There are several different factors relevant to our sensitivity analysis:

1. Distribution of the path priors across the tree, e.g., a path may have an extremely low prior probability vs. one of approximately uniform path priors.

2. The distribution of the cues across the events, e.g., multiple cues assigned to one event versus a single cue assigned to another event.

3. Cue strength variation, e.g., we can choose to vary the strength all cues of one qualitative strength assessment (e.g. “strong” or “weak” cue), choose to change the class of various cues, or individually adjust the strengths sof the cues.

4. Tree structure, i.e., some branches of the tree may have considerably more associated events than other branches.

5. Cue observation pattern, e.g., randomly scattered cues versus cues aligned along a scenario path.

While not comprehensive, the sensitivity analysis performed here is meant to span or generalize results for a range of these confounding parameters.

5.1. Sensitivity to uncertainty in cue strength of a single cue

The first case we analyze is the very simple case of having one cue present in one location of the tree. The results from this study demonstrate the first order sensitivity of cues. These general results of sensitivity extend to different cases of cue observation, for example having a single cue appear in multiple places in the tree.

In order to study the first order effects of a cue strength we simplify the example tree that was shown in Figures 3 and 4. We consider the case where each branch has equal priors; wherever the tree splits into two, the prior probability is split evenly. We also remove all cues except for one hypothetical cue assigned to the state corresponding to *Weapon fabrication in Country A* after *Sea transit to Fabrication*. While such a cue doesn't exist in Table 3, a possible cue could be "local observation of cargo unloading at Country A seaport." In doing this, all other possible confounding factors are removed and thus we can clearly see the first order effects of cue value on the posterior probabilities compared to the priors.

In examining the sensitivity for varying cue strength values there are two indicators that are useful. First we look at the difference between the posterior and prior probabilities of the event, Δ. This gives the absolute sense of how much things change. We also examine the absolute value of instantaneous slope (or derivative) of the Δ. The higher the slope of the Δ, the more sensitive a cue is to the cue strength value around that point. This is useful in giving a general idea of sensitivity of a cue at a given cue strength value.

Recall the value of $\frac{P_D}{P_{FA}}$ plays the key role in affecting the posterior when a cue is observed. We examine the path probability up to the event with the observed cue for a range of $\frac{P_D}{P_{FA}}$ values with a fixed $P_{FA} = 0.1$ in Figure 7. Looking at the absolute value of the slope of the Δ over the prior, we see that at lower values of $\frac{P_D}{P_{FA}}$, the slope is higher than at higher values of $\frac{P_D}{P_{FA}}$. This indicates that the posterior is more sensitive to changes in the value of weaker cues than stronger cues. However, the absolute difference in the posterior changes are still small in the examples above.

Intuitively, the reason why the result is less sensitive to stronger cues is because as a path's posterior probability approaches the maximum value of 1 with an increasingly stronger cue, there can be less impact that that cue can have on the posterior probability. Thus the sensitivity approaches the limit asymptotically. That is, the difference between a strong cue and a very strong cue is less than the difference between a very weak cue and a weak cue, and this is reflected in the posterior.

Additionally, we examine the effect not observing a cue as a function of P_D in Figure 8. Looking at the absolute value of the slope of the Δ, we can see the result of selecting value of a "low likelihood cue" is less sensitive to that of selecting the value of a "high likelihood cue." That is, the higher the P_D chosen, the more sensitive the posterior is to perturbations. Also, the value of

Figure 7 **Sensitivity when observing a cue for various values of cue strength as defined by $\frac{P_D}{P_{FA}}$. We fix value of $P_{FA} = 0.1$. In the interest of completeness, we extend the test to values below $\frac{P_D}{P_{FA}} < 1$. Top: Plot of the difference (Δ) between the posterior and prior probabilities of the event with the observed cue. Bottom: Absolute value of the instantaneous slope of the Δ between the posterior and prior probabilities.**

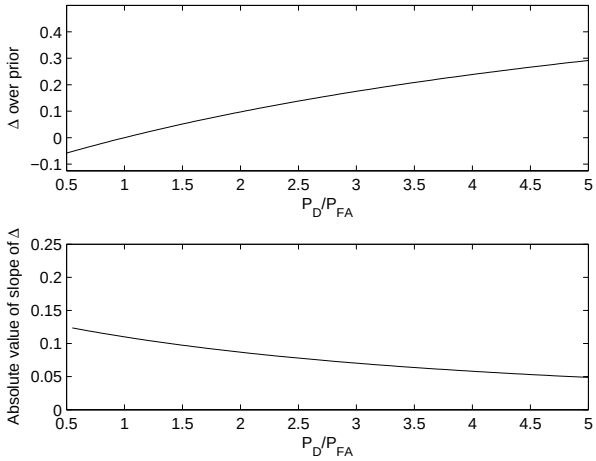
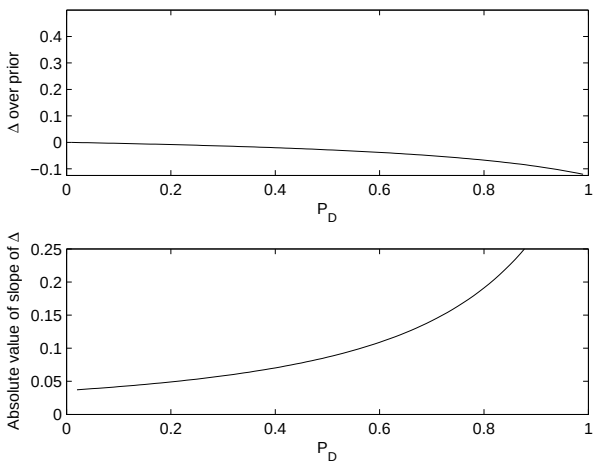


Figure 8 **Sensitivity when not observing a cue for various values of cue likelihood as defined by P_D. The $\frac{P_D}{P_{FA}} = 2$ for all values of P_D. Top: Plot of the difference (Δ) between the posterior and prior probabilities of the event with the non-observed cue. Bottom: Absolute value of the instantaneous slope of the Δ between the posterior and prior probabilities.**



the Δ over the prior for lower likelihood cues show that not-observing a cue will have very little impact on the posterior probabilities.

The effect on the posterior of not observing a cue increases as P_D approaches 1 derives from the fact that at a $P_D = 1$, the posterior probability must equal 0. Thus the slope of the Δ increases as P_D approaches 1. Intuitively, this means that the more certain you are of seeing a cue given an

event occurring, the more likely you are to believe that the event did not occur if you do not see the cue.

Looking at the effect of cue value assignments on upstream events the same overall curve shapes as Figures 7 and 8 also hold. The only difference is that the slope of Δ for the upstream events is generally lower. That is, there is not as large of an impact on higher level upstream event priors when selecting cue values associated with a downstream event. This is in line with the intuition that higher level events have more paths through the event tree to distribute the probability mass.

In summary, the first order effects show that the relative impact of the uncertainty when assigning values to $\frac{P_D}{P_{FA}}$ for an observed "weak cue" is slightly larger than the impact when assigning values for stronger cues. However, the absolute impact overall for assigning values to the $\frac{P_D}{P_{FA}}$ categories is still small, especially on higher level events. For a non-observed "low likelihood cue," the relative impact on the posterior of assigning an uncertain value is less than for higher likelihood cues.

Additionally, we also observe that if a weak cue is misclassified as a strong cue (or vice versa), then the absolute difference compared to the prior can be significant. This difference is expected since model cannot correct for such misclassifications.

These results also hold for other cases spanning the parameter space. In general the assignment of the Bayes' factor has a higher impact on the posterior at low values than at high, and in general the value of the likelihood is more influential at high values than low.

We also extended this simple case of one single cue in one location on the event tree, we look at how having the single cue appear in more than one location on the event tree. The relative trends remain the same as the case when the cue exists for only one event in the tree. That is, assigning a lower value of $\frac{P_D}{P_{FA}}$ is more sensitive to uncertainty than selecting a higher $\frac{P_D}{P_{FA}}$. And, selecting a lower value of P_D is less sensitive to uncertainty than higher values of P_D.

5.2. Sensitivity to uncertainty of cue strength of entire class of cues

We have observed how a single cue is sensitive to the value of its strength; we now look to see how changing the cue strength across an entire class can affect inference.

We return to the original example tree of Section 4 with all of the multiple cues scattered across the tree in a mostly random manner as defined in Section 4.3. We examine this case becasue this is the case where we expect the increasing cue strength not to make a large difference.

We now adjust the P_D for all "weak cues" that originally have a value of $P_D = 0.1$ and $P_{FA} = 0.08$. Other cues with different P_D and P_{FA} are left unadjusted from their values in Table 4. This does also affect cues that are not observed and so there will also be negative evidence affecting the inference. The cues configuration remains the same as in Section 4.3.

To see the effect changing the cue value assignment of a weak cue under this condition we examine three alternative values of $\frac{P_D}{P_{FA}}$ by changing the P_{FA} so as to reduce the impact of not observing

Table 11 Posterior values for various cue strength values when cue observations are scattered across the event tree.

Event	Prior	$\frac{P_D}{P_{FA}} = 1.11$	$\frac{P_D}{P_{FA}} = 1.25$	$\frac{P_D}{P_{FA}} = 2.5$	$\frac{P_D}{P_{FA}} = 5$
Steal Material	0.40	0.44	0.43	0.38	0.36
Steal Weapon	0.60	0.56	0.57	0.62	0.64
Sea Transit Material to Fabrication Country	0.24	0.26	0.25	0.22	0.21
Air Transit Weapon To US	0.36	0.34	0.34	0.37	0.39
IND Detonation	0.01	0.01	0.01	0.01	0.00
Weapon Detonation	0.04	0.03	0.03	0.02	0.01

cues. First we double and quadruple the strength of a weak cue. Then we weaken the effect of the likelihood by splitting the difference between the original P_D and P_{FA} effectively weakening the cue by half, i.e. $P_{FA} = 0.09$. In each case the P_D remains 0.1. Thus only the key $\frac{P_D}{P_{FA}}$ strength metric is changing. We compare these new cue strength values to the original values in Table 8 of Section 4.3.

In Table 11, we see that the doubling or quadrupling makes small changes to the posterior probability relative to the prior probabilities. Even though the cues are still relatively well spread apart, there is still a pattern on the steal weapon side since this tree is small. This pattern is highlighted by the shifting of a small amount of probability mass to the weapon side of the tree as the cue strength increases.

6. Implementation and Scalability

The model presented in this report has been demonstrated with a small example. To be useful, the model must be scalable to much larger problems and address implementation challenges associated with generating and ingesting required data.

The model described in this report presupposes access to the defined cues in a structured way. That is, we assume the cues are both available and in a format that the model can easily utilize. This assumption implies at least three steps have been taken regarding the collection of intelligence data prior to using the model:

1. Gaining access to all of the necessary sources of information.
2. Converting raw intelligence (e.g., free text, video, image, or audio data) into structured data that is easily interpretable by a machine or human analyst.
3. Converting structured data into cues relevant to the model.

Additionally, an important step in the implementation of the model is to elicit judgments from experts on the nuclear terrorism scenarios, the cues, and the cue strengths. Specifically, the model relies on estimates of the probability of observing the relevant cues when a scenario is underway, as well as observing the cues when the scenario is not underway. The latter can alternatively

be ascertained from historical data, although consistency with the expert judgments must be maintained.

For demonstration purposes, the examples in this report have been based on a very small set of scenarios with a corresponding small event tree. To be operationally useful, the event tree for nuclear terrorism would need to be significantly larger than those shown above. The algorithms used in the model run in $O(n)$ time where n is the number of nodes in the tree, meaning that the computational burden grows in proportion with the number nodes in the tree. It should be noted that the size of the tree, n, can increase rapidly with the number of events in the model. However, the computations performed at each iteration are very simple. Therefore, the model is scalable and can accommodate very large event trees without a heavy computational burden.

7. Conclusions and Future Work

We have demonstrated a method for exploiting ambiguous cues given some prior belief on adversary behavior (derived from a decision analysis framework or expert assessment). The strength of this model is that it can analyze a large number of ambiguous cues in the context of a large number of possible nuclear threat scenarios. The model provides a platform to combine information in the scenario tree provided by nuclear weapons experts with information and judgments provided by intelligence analysts. The analysis provided by this tool can be used to inform decisions on how to focus resources on events earlier in the scenario (upstream) relative to the current time, or it can be used to inform decisions about subsequent (downstream) countermeasure deployment decisions.

Our method is based on Bayes' Rule to incorporate cues to update prior probabilities. We applied this methodology to a small hypothetical example and our method is scalable to larger trees and more cues. We also studied the sensitivity of the methodology to the cue strength values and scenario modeling assumptions. We see that all else being equal, the pattern of observed ambiguous cues within the scenario tree can have a significant impact on the resulting assessment of adversary state. That is, many ambiguous cues observed along a single scenario path provide powerful evidence to support the likelihood of that path; alternatively, many ambiguous cues observed scattered across the event tree do not significantly change the prior belief about the adversary.

To address the uncertainties in the scores, we performed an initial sensitivity analysis on the score value. When a cue is observed, the model is more sensitive to uncertainties in the $\frac{P_D}{P_{FA}}$ strength of weaker cues than of stronger cues. A cue with a lower P_D is less sensitive to small changes in its value than a cue with a higher P_D.

A number of challenges remain to implement the methodology proposed in this report:

- This methodology considers the relative likelihood of attack scenarios, and doesn't consider the case where no threat attack is underway. Enhancements to the model could include the ability to shed light on the likelihood a threat scenario is underway.
- The methodology can account for correlation, but it requires direct elicitation of correlations from subject matter experts. Examples of correlated cues were not shown here. The probability of observing a cue depends only on the corresponding adversary event and not on the path taken to reach the event. However, cues can be correlated with each other in many ways. When correlation among observed cues exists due to non-threat activity (i.e., benign activity that generates a series of cues that could be mistaken for threat activities), resources might be diverted away from potential threat scenarios to adjudicate the false alarm. Also, correlation among cues generated by threat activities can be more important to model. For example, consider two cues that are individually ambiguously indicative of a threat scenario underway, but when observed together are more indicative than their combined affect would otherwise yield. This combination of cues can be a powerful indicator of threat activity only if the correlation among the cues is properly treated. Capturing this type of correlation involves additional expert elicitation or historical data analysis.
- A closely related issue is inclusion of known structural dependencies of cues. For example, two cues indicating the locations of a threat at two points in time may be inconsistent with the time required to travel between the two points. In this case, we can infer that the two cues cannot both be associated with a single threat scenario. Incorporating this type of information would increase the fidelity of this model.

Acknowledgments

This work was funded by the Domestic Nuclear Detection Office at the Department of Homeland Security. Prepared by LLNL under Contract DE-AC52-07NA27344. Document Number LLNL-JRNL-554944. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or Lawrence Livermore National Security, LLC, and shall not be used for advertising or product endorsement purposes.

References

Army, U.S. 2007. *A Military Guide to Terrorism in the Twenty-First Century*. US Army Training and Doctrine Command.

Bier, V. M. 2007. Choosing what to protect. *Risk Analysis* **27** 607–620.

Bunn, M., A. Wier, J. Holdren. 2003. Controlling nuclear warheads and materials: A report card and action plan. Tech. rep., Nuclear Threat Initiative (NTI).

Chengara, A. 2004. Indicators of terrorist attacks. *Bharat Rakshak - Security Research Review* **1**(1).

Hausken, K., J. Zhuang. 2011. Governments' and terrorists' defense and attack in a t-period game. *Decision Analysis* **8** 46–70.

Heckerman, David. 1995. A tutorial on learning with bayesian networks. Tech. Rep. MSR-TR-95-06, Microsoft Research.

Kass, R.E., A. E. Raftery. 1995. Bayes factors. *Journal of the American Statistical Association* **90**(430) 773–794.

Khalsa, S. 2004. *Forecasting Terrorism: Indicators and Proven Analytic Techniques*. The Scarecrow Press.

Krupka, J.J. 2006. Overcoming ambiguity at the operational level. Tech. Rep. ADA463433, Naval War College.

Maurer, S. M. 2009. *WMD Terrorism: Science and Policy Choices*. MIT Press.

Rabiner, L.R. 1989. A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE* **77**(2) 257–286.

Rabiner, L.R., B. Juang. 1986. An introduction to hidden markov models. *ASSP Magazine, IEEE* **3**(1) 4–16.

SCN, Secure Community Network. 2005. Seven signs of terrorist activity (sourced from the michigan state police). [Online]. URL `ttp://www.scnus.org/page.aspx?id=101218`.

Smith, A. 2005. Detecting terrorist activity: Defining the states threshold of pain. *Australian defence force journal* 30–44.

Smith, B. L., K. R. Damphousse, P. Roberts. 2006. Pre-incident indicators of terrorist incidents: The identification of behavioral, geographic, and temporal patterns of preparatory conduct. Tech. Rep. 214217, U.S. Department of Justice.

Tu, H., J. Allanach, S. Singh, K. R. Pattipati, P. Willett. 2004. The adaptive safety analysis and monitoring system. *SPIE Defense and Security Symposium*. 12–16.

Wyllie, D. 2011. 8 pre-attack indicators of terrorist activity. [Online]. URL `http://www.policeone.com/federal-law-enforcement/articles/3421116-8-pre-attack-indicators-of-terrorist-activity/`.