A Framework for Estimating the Value of Deterrence

Richard S. John[a], Robin Dillon[b], William Burns[c], and Nicholas Scurich[d]
[a] University of Southern California, Los Angeles, California, USA, richardj@usc.edu
[b] Georgetown University, Washington, D.C., USA, Robin.DillonMerrill@georgetown.edu
[c] California State University, San Marcos, California, USA, bburns@csusm.edu
[d] University of California, Irvine, California, USA, nscurich@uci.edu

Abstract. This paper presents a framework for calculating the value of deterrence related to countermeasures implemented to mitigate an attack by an adaptive adversary. We offer a methodology for adapting Defender-Attacker Decision Trees to partition the utility of countermeasures into three components: (1) threat reduction (deterrence), (2) vulnerability reduction, and (3) consequence mitigation. The Expected Utility of Imperfect Control (EUIC) attributable to a specific implementation of the countermeasure is based on calculations from decision analysis and is defined as the difference in the expected utilities of the no countermeasure branch and the branch representing the countermeasure variant. The EUIC represents the net benefit of implementing the countermeasure, including all costs associated with development, implementation, and operation. Benefits primarily derive from three sources: (1) changes in attack probability (threat reduction), (2) changes in detection probability (vulnerability reduction), and (3) changes in the distribution of attack outcomes (consequence mitigation). We partition the EUIC and estimate the unique portion attributable to threat reduction, vulnerability reduction, and consequence mitigation. Calculations follow a subtraction logic, similar to those used to calculate the Value of Information (VOI). We provide example applications of the Value of Deterrence in an airport security domain. The proposed framework provides a methodology for explicitly accounting for deterrence in benefit-cost analyses.

1. INTRODUCTION

Since 9/11, decision and risk analysis tools have been used in many contexts to help improve risk assessment and decision support related to terrorism threats. Probabilistic Risk Analysis (PRA) was proposed as a methodology to assess risks from adaptive adversaries [1, 2]. Defender decisions related to resource allocations and countermeasure deployment utilized PRA as part of an adversary decision analysis incorporating Defender-Attacker decision trees in the context of bioterrorism [3], Man-portable air-defense systems (MANPADS) terrorist threats to commercial air travel [4, 5], and other terrorist threats [6]. A project management framework was utilized to model attacker uncertainties related to terrorist dirty bomb threats to a port [7]. Value-focused thinking [8] has been used to structure the values of terrorist groups [9, 10], and multiattribute utility models [11] have been applied to quantify defender uncertainties about terrorists' preferences and trade-offs [12]. Defender-Attacker games that explicitly account for deterrence have been used to inform defender resource allocation decisions [13-25]. These studies have generally been well cited in the literature, an academic measure of success, but when risk assessment and decision support are working best in this context, terrorist activities are deterred. It has generally been difficult to quantify the benefits of some future uncertain event not happening (being deterred). This paper contributes to the literature by providing a methodology for quantifying deterrence based on partitioning the benefits of different countermeasures.

While the methodologies and applications cited above account for threat reduction and deterrence, none provides an explicit procedure for partitioning the benefits of countermeasure alternatives. This is an important concern given that countermeasures are often presumed to have deterrent effects that are difficult to characterize and quantify. Threat reduction and deterrence are critical aspects of most countermeasures, and it is important to quantify their impacts on the defenders' expected utility. As demonstrated in this paper, benefits related to deterrence are often confounded with other countermeasure benefits, such as reduction in both vulnerability and mitigation of the consequences of a successful attack. We utilize Defender-Attacker decision trees to partition countermeasure benefits

and estimate the unique contributions of the countermeasure to reduce threats, vulnerability, and consequences. Alternative countermeasures can thus be characterized in terms of a profile of expected reduction in threat, vulnerability, and consequences.

2. PARTITIONING COUNTERMEASURE BENEFITS

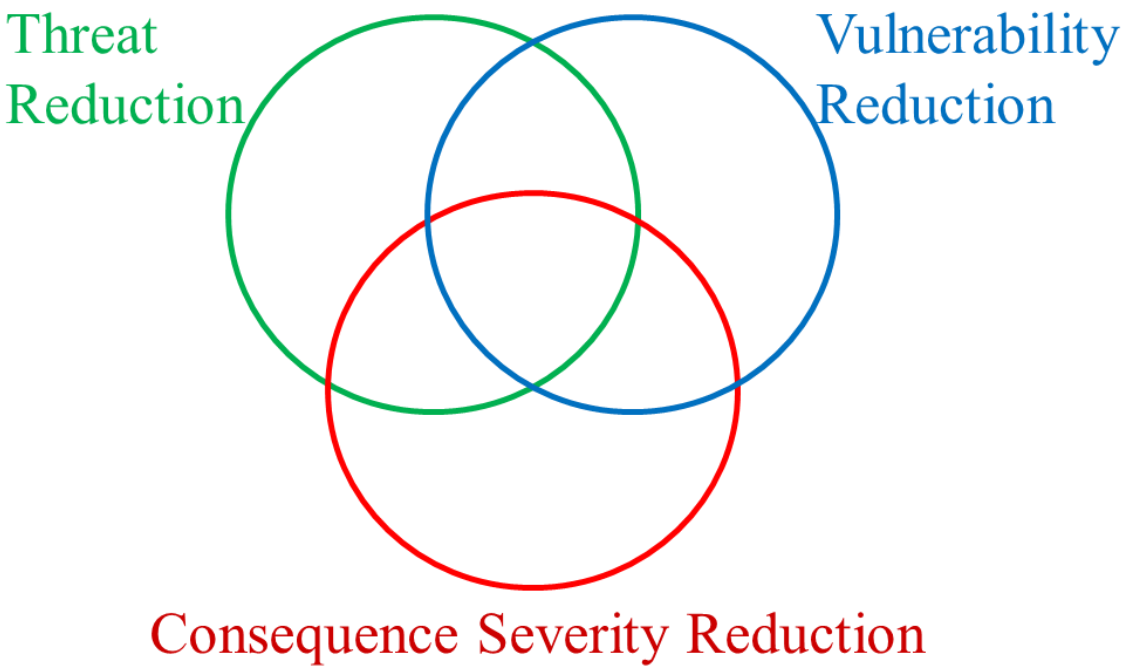
We describe a methodology to adapt Defender-Attacker Decision Trees (DADTs) to facilitate partitioning the expected utility of countermeasures into three components: threat reduction (deterrence), vulnerability reduction, and consequence mitigation. The DADT analysis compares two or more branches from a defender decision node that includes at least two alternatives: implement the countermeasure and do not implement the countermeasure (status quo). Variations on the nature and scope of the countermeasure implementation could also be included with additional alternatives represented by additional branches from a defender decision node. The Expected Utility of Imperfect Control (EUIC) attributable to a specific implementation of the countermeasure is based on standard calculations from decision analysis and is defined as the difference in the expected utilities of the no countermeasure branch and the branch representing the countermeasure variant [26, 27]. The EUIC represents the net benefit of implementing the countermeasure, including all costs associated with development, implementation, and operation. Benefits would largely derive from three sources: (1) changes in attack probability (threat reduction), (2) changes in detection probability (vulnerability reduction), and (3) changes in the distribution of attack outcomes (consequence mitigation).

In order to partition the EUIC and estimate that unique portion attributable to deterrence, the countermeasure branch must be modified such that there is no change in attack probabilities and hence no deterrence effect. The difference between the expected utilities of the countermeasure implementation branch with and without a change in attack probabilities represents the Expected Utility of Imperfect Deterrence (EUID). Note that this calculation removes the benefit of the countermeasure for vulnerability reduction and consequence mitigation since both are considered. This subtraction logic is similar to that used in Value of Imperfect Information calculations [28] and allows for the calculation of the part of the countermeasure net benefit that can be attributable to a reduction in threat through a reduction in attack probability.

The Expected Utility of Perfect Deterrence (EUPD) can then be calculated by assuming the attack probability goes to zero and the threat is eliminated. Again, the subtraction strategy can be applied to compare the expected utilities of the countermeasure branch with no change in attack probability and a zero-attack probability. This calculation is useful for comparing the EUID estimate to determine the relative deterrent effect achieved by the countermeasure.

Similar calculations can also be made to obtain estimates for the utility of the countermeasure in terms of both the value of vulnerability reduction (VoVR) and the value of consequence reduction (VoCR). By comparing the countermeasure branch with and without changes in detection probability or changes in outcome probabilities following a completed attack, estimates of the countermeasure's reduction in vulnerability and consequence reduction can be calculated and compared to the reduction in threat (deterrence effect). Note that the sum of the three components of imperfect control (threat reduction, vulnerability reduction, and consequence mitigation) will not necessarily be additive. That is, the total EUIC will not, in general, be the sum of the three component utilities. Figure 1 presents a Venn diagram illustrating the overlap in the expected benefits of reduction in threat (deterrence), vulnerability, and severity of consequences of a successful attack. As indicated in Figure 1, any two countermeasure benefit components may overlap, and a three-way overlap is also possible. The unique contribution of each component is that portion that does not overlap either or both of the other two. This result is similar to the well-known result in expected utility of information (EUOI) analyses for multiple sources of information.

Figure 1. Venn Diagram of the Potential Overlap in Three Countermeasure Benefit Components

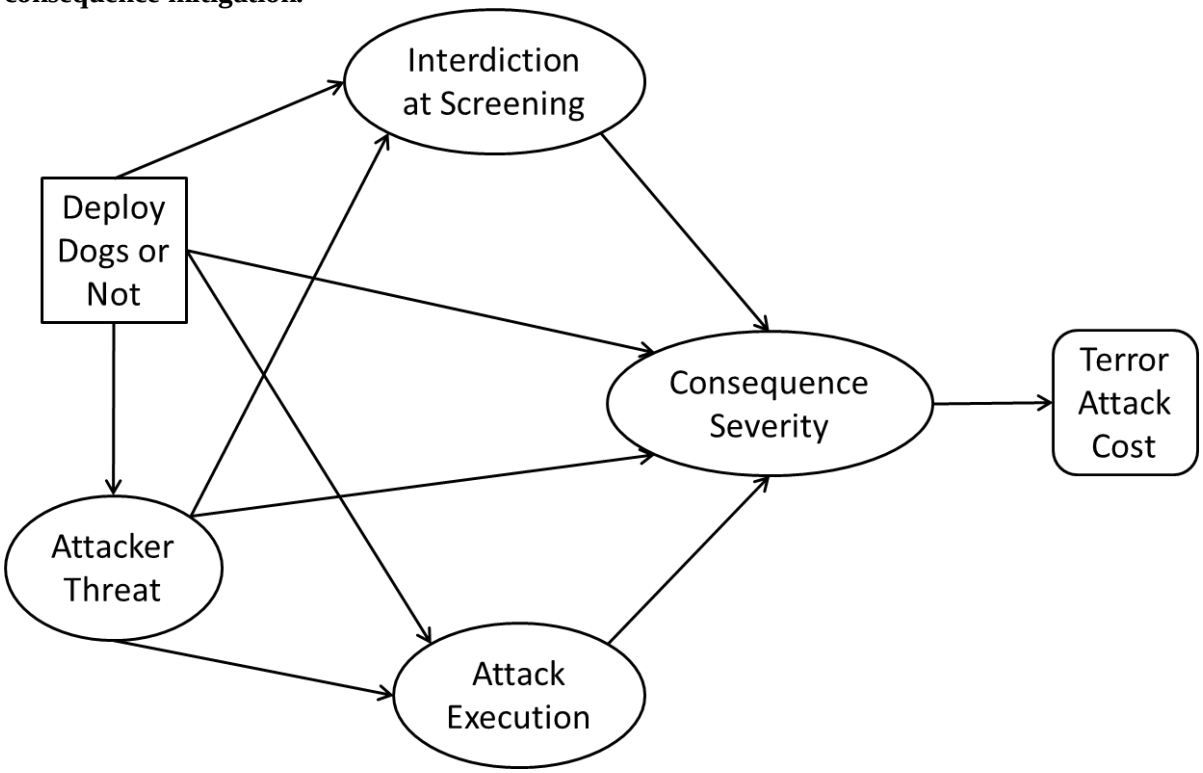


3. COUNTERMEASURE EXAMPLE

3.1. TSA Canine Deployment at Airports

We will illustrate some of these calculations with a stylized example of a decision by the TSA to deploy canines as a countermeasure at a particular airport, over and above other countermeasures already deployed at airport passenger screening checkpoints. An example DADT is presented in Figure 1 in the compact form of an Influence Diagram representing a single defender's decision to deploy dogs or not and four different uncertainties dependent on that decision, including the attacker threat (explosives attack, a firearm attack, vs. no attack), two vulnerability uncertainties (interdiction by detection of the weapon, or not, and successful execution of the attack, or not), and the severity of the consequences (extreme, moderate, or minimal). Note that arcs represent probabilistic dependency only and not necessarily a causal relationship. The example results presented below were computed using a detailed DADT representation, including notional values for probabilities and costs of outcomes, not including the fixed cost of deploying the dog countermeasure. Note that probability distributions and outcome costs are notional; an actual EUIC analysis would require inputs from SMEs, with or without sensitive information. It is important to understand that the dogs have multiple influences on the expected cost of a terror attack on a commercial plane via the probability distributions for the uncertainties, conditional on countermeasure deployment. In most cases, adding another countermeasure has a decreasing marginal benefit to a portfolio of countermeasures (i.e., "ladling on can only do so much") [1]. The influence diagram is translated into an equivalent defender-attacker decision tree below to illustrate the decomposition of benefits into the three categories previously described.

Figure 2. Influence Diagram representation of decision to deploy dogs as countermeasures at airport passenger security checkpoints, including deterrence, vulnerability reduction, and consequence mitigation.



3.2 Partitioning Annual Expected Benefits of Deploying Canines at Airport

In estimating the economic costs associated with attacks and countermeasures, this example only considers damage from the attack. To keep the example simple, we did not include the costs of the countermeasure, but in realistic case studies, we would include such costs. Additionally, we assume four possible attack outcomes: a failed attack with $0 costs, a minimal damage attack with $10 million in costs, a moderate damage attack with $1 billion in costs, and a maximum damage attack with $10 billion in costs. These are consistent within an order of magnitude of consequence estimates that include indirect costs associated with terrorist attacks on aircraft [29]. We use notional probabilities and consider the threat annually (i.e., attack in the next year).

Deterrence is achieved by reductions in the probability that the attacker will choose to attack using an explosive device that the canines, as a countermeasure, have a high probability of detecting. Vulnerability reduction is achieved by increases in the likelihood of detection and interdiction of weapons at the checkpoint, particularly explosive devices. Consequence reduction is achieved if the possible destruction scenarios change, for example, because of less effective IEDs being created to attempt to avoid detection. In this example (again for simplicity), we do not consider consequence reductions but instead examine only deterrence and vulnerability reduction. Quantifying the reductions from the countermeasures and incorporating them into a structured decision tree approach allows us to quantify the expected utility of perfect and imperfect deterrence.

Figure 3 shows the portion of the decision tree for the case where TSA would decide not to deploy canine units at Airport X. The probabilities are notional but, in theory, would include the other countermeasures available at the Airport screening location. Figure 4 shows the portion of the decision tree where TSA would deploy canine units. As shown in Figures 3 and 4, the canine countermeasure is imperfect deterrence because while it shifts the terrorist's attack probabilities, the probability of attack in the next year with the countermeasure deployed is still not zero.

Figure 3. Branch of Defender-Attacker Decision Tree following Decision not to Deploy Canine Units (consequences in $M)

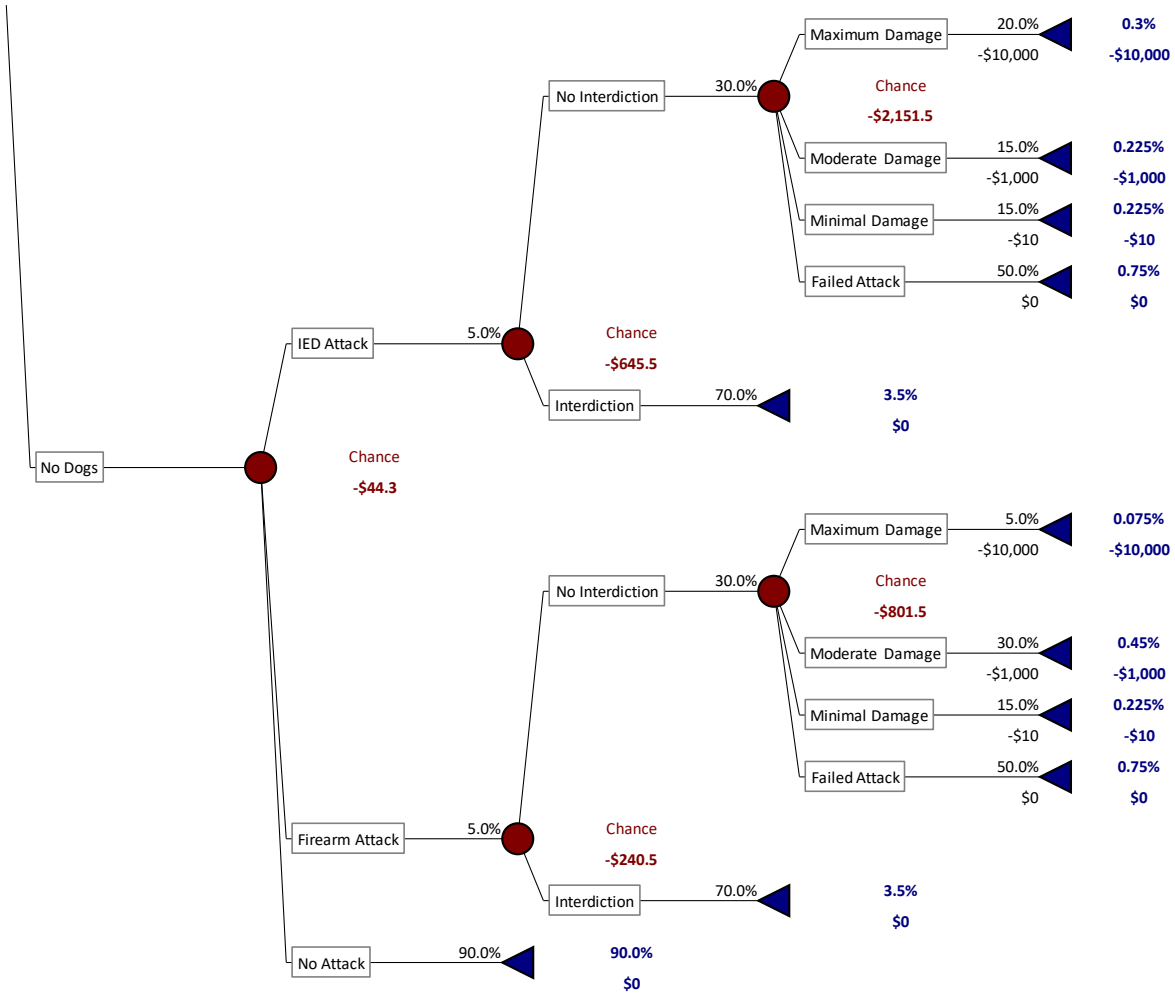
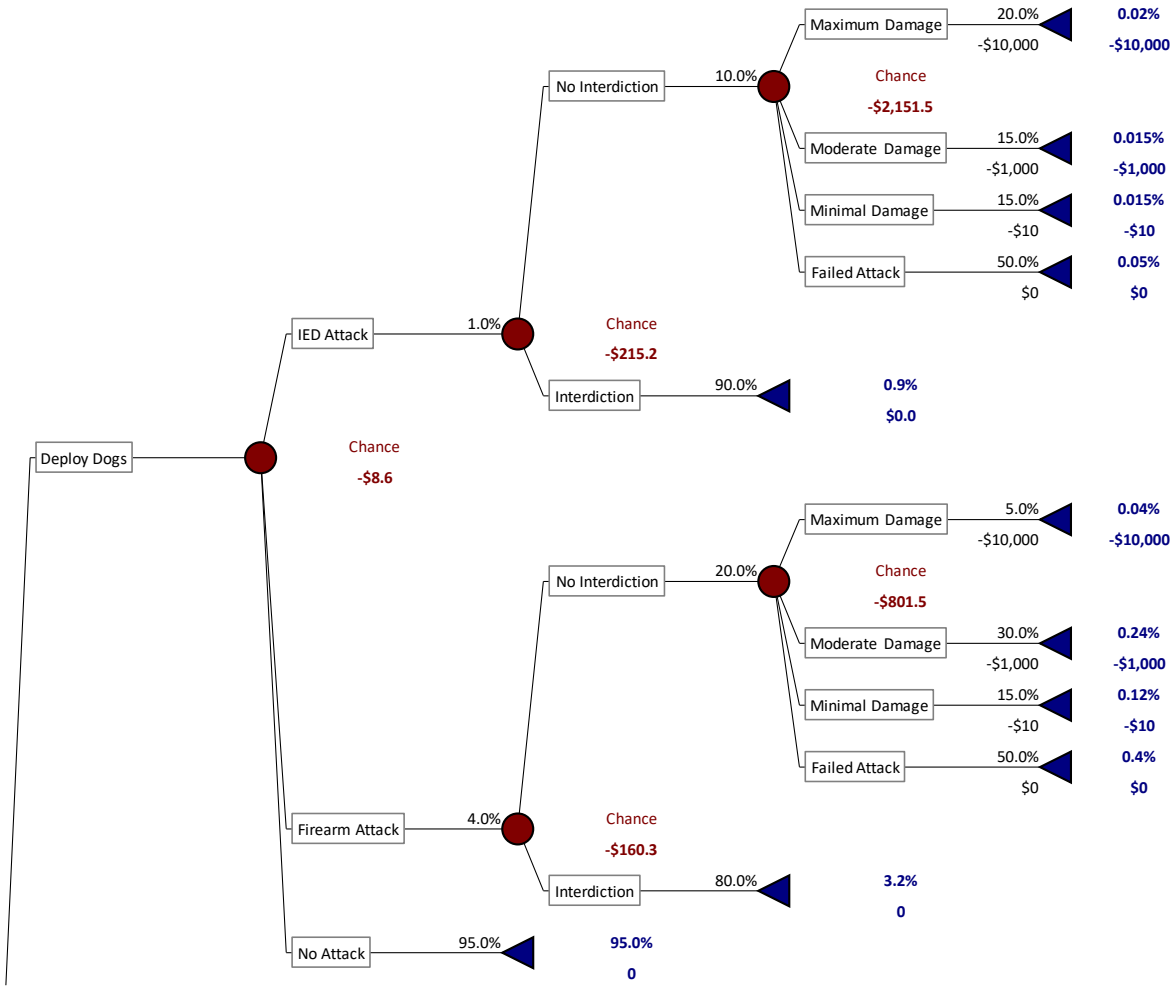


Figure 4. Branch of Defender-Attacker Decision Tree following Decision to Deploy Canine Units (consequences in $M)



Comparing Figures 3 and 4, the probabilities of attack in the next year change between the two branches, and the probabilities of interdiction, given an attack, change. As noted above, the consequences given an attack do not change in the example, so there is no expected consequence reduction. In this example, the attacker is less likely to attempt the IED attack, more likely to attempt a firearm attack, and more likely to be deterred (value of deterrence). The TSA is also more likely to interdict the attack for both IEDs and Firearms (value of vulnerability reduction). Thus, assuming these are representative probabilities and consequences, the economic benefit of the countermeasure in terms of consequences avoided is $35.7M (i.e., -$8.6M – (-$44.3M)), the difference between the expected outcome if the TSA does and does not deploy the canine countermeasure.

We further explore isolating the value of deterrence and vulnerability reduction by comparing the countermeasure branch with and without changes in attack strategy or detection probability to get estimates of the individual components: the countermeasure's reduction in vulnerability and reduction in threat (deterrence effect). Note that the sum of the components is not necessarily additive.

Examining Figures 3 and 4, in determining EUIC for this example, the deployment of the countermeasure changed the expected probabilities of the terrorist's attack choice (including no attack) and the expected probabilities of the TSA interdicting the attack. In order to isolate the benefits of deterrence only, the decision tree is recalculated with the interdiction probabilities for all attacks at 30% (the same as Figure 3). If interdiction does not change and only the probabilities of the terrorist's attack

choice change, then the expected utility of no countermeasure deployment is -$44.3M and of deploying the canine countermeasure is -$16.1M. Thus, the calculated EUID is $28.2M.

If there is perfect deterrence from the countermeasure, such that the attacker chooses to not attack with a 100% probability, then estimating this change in Figures 3 and 4, the expected utility of the canine countermeasure will be $0 (no attack losses and again we are not considering costs in this example). Thus, the EUPD is $44.3M (the expected losses in the no countermeasure deployment option).

To determine the benefit of the countermeasure in terms of vulnerability reduction, the expected probabilities of the attack method are unchanged between Figures 3 and 4, and only the interdiction probabilities change. If the attack method does not change and only the probabilities of interdiction change with or without the countermeasure deployment, then the expected utility of no countermeasure deployment is still -$44.3M and of deploying the canine countermeasure is -$18.8M for a Value of Vulnerability Reduction of $25.5 M.

As discussed above, in some situations, the countermeasure could also change the expected consequences or the probability of different consequences. While this possibility is not included in this example, a similar process would be used to determine the benefit of the countermeasure in terms of consequence mitigation, the expected probabilities of the attack method and the probabilities of interdiction would be held constant, and the two branches of the decision tree would be recalculated to determine the impact of the countermeasure on the expected consequences.

Table 1 - Summary Table of Calculations for the Canine Countermeasure Example

Expected Utility of Imperfect Control (EUIC)	The net benefit of implementing the countermeasure	The difference between the expected utilities of the no countermeasure branch and the countermeasure variant: -$8.6M – (-$44.3M) = $35.7M
Expected Utility of Imperfect Deterrence (EUID)	The net benefit of deterrence from the countermeasure	The difference between the expected utilities of the no countermeasure and the countermeasure branches isolating only the changes in attack probabilities: -$16.1M – (-$44.3M) = $28.2M
Expected Utility of Perfect Deterrence (EUPD)	The net benefit if the countermeasure completely deters the attacker	The countermeasure is 100% effective at deterring the attack so that it does not happen: $0 – (-$44.3M) = $44.3M
Value of Vulnerability Reduction (VoVR)	The net benefit of improved interdiction from the countermeasure	The difference between the expected utilities of the no countermeasure and the countermeasure branches isolating only the changes in interdiction probabilities: -$18.8M – (-$44.3M) = $25.5M
Value of Consequence Reduction (VoCR)	The net benefit of reduced consequences from the countermeasure	The difference between the expected utilities of the no countermeasure and the countermeasure branches isolating only the changes in consequences is *not applicable in this example*

While the framework described here is based on decision and risk analysis and does not require the development of new theoretical tools, the implementation of such a framework for the complexities of a real example is numerous. As mentioned, this example ignored: 1) the costs of the countermeasures, 2) the interaction among countermeasures for detection, 3) the decreasing marginal detection benefit of additional countermeasures depending on the current portfolio, and 4) the possibility for consequence reduction from countermeasures.

4. CONCLUSION

We have presented a framework adapted from Defender-Attacker decision trees to partition and quantify expected countermeasure benefits related to reductions in threat, target vulnerability, and consequence severity. We have provided a proof-of-concept example of an analysis of the benefits of deploying canines to a particular airport to reduce terrorist threat, target vulnerability, and consequence severity.

Both the probabilities and consequence estimates in the decision tree represent the uncertainties and values of the defender. In most cases, probabilities related to target vulnerability and consequence severity can be modeled using data related to the countermeasures and target studied. In contrast, changes in threat probabilities related to the introduction of alternative countermeasures require the defender to utilize a theory of mind of the attacker. That is, the defender must have some idea of how the attacker would perceive the various countermeasures under consideration and decide to either attack or not. Thus, accurate quantitative estimates of the value of deterrence require an understanding of attacker preferences, including the perception of uncertainty, risk attitude, and trade-offs among conflicting objectives. In addition, heuristics and biases in the attackers' judgment and decision making will also impact the attacker's decision to attack or not.

Accurate estimates of the value of deterrence will require further development of a psychology of deterrence that accounts for extra-rational aspects of attacker perceptions and preferences that may have a substantial impact on an adaptive adversary's decision to attack or not. Some recent progress has been made on the psychology of deterrence in the domain of airport security checkpoints [30], cybersecurity [31,32], and other security contexts [33, 34]. We anticipate that applications of the partitioning methodology to identify and quantify unique contributions of countermeasures in terms of threat reduction, target vulnerability reduction, and consequence severity mitigation will require further advances in the psychology of deterrence.

References

[1] E. Paté-Cornell and S. Guikema. *"Probabilistic modeling of terrorist threats: A systems analysis approach to setting priorities among countermeasures",* Military Operations Research, 7, pp. 5-23, (2011).

[2] B. C. Ezell, S. P. Bennett, D. v. Winterfeldt, J. Sokolowski, and A. J. Collins. *"Probabilistic risk analysis and terrorism risk",* Risk Analysis, 30, pp. 575-589, (2010).

[3] G. S. Parnell, C. M. Smith, and F. I. Moxley. *"Intelligent adversary risk analysis: A bioterrorism risk management model",* Risk Analysis, 30, pp. 32-48, (2010).

[4] D. v. Winterfeldt and T. M. O'Sullivan. *"Should we protect commercial airplanes against surface-to-air missile attacks by terrorists?"* Decision Analysis, 3, pp. 63-75, (2006).

[5] R. J. B. Garcia and D. v. Winterfeldt. *"Defender–attacker decision tree analysis to combat terrorism",* Risk Analysis, 36, pp. 2258-2271, (2016).

[6] J. Merrick and G. S. Parnell. *"A comparative analysis of PRA and intelligent adversary methods for counterterrorism risk management",* Risk Analysis, 31, pp. 1488-1510, (2011).

[7] H. Rosoff and D. v. Winterfeldt. *"A risk and economic analysis of dirty bomb attacks on the ports of Los Angeles and Long Beach",* Risk Analysis, 27, pp. 533-546, (2007).

[8] R. L. Keeney. "*Value-focused thinking: A path to creative decisionmaking,*" Harvard University Press, (1996), Cambridge.

[9] G. L. Keeney and D. v. Winterfeldt. *"Identifying and structuring the objectives of terrorists"*, Risk Analysis, 30, pp. 1803-1816, (2010).

[10] J. Siebert, D. v. Winterfeldt, and R. S. John. *"Identifying and structuring the objectives of the Islamic State of Iraq and the Levant (ISIL) and its followers"*, Decision Analysis, 13, pp. 26-50, (2016).

[11] R. L. Keeney and H. Raiffa. *"Decisions with multiple objectives: Preferences and value trade-Offs,"* Wiley, (1976), New York.

[12] H. Rosoff and R. S. John. *"Decision analysis by proxy for the adaptive adversary,"* In A. Abbas, M. Tambe, & D. von Winterfeldt (Eds.), Improving homeland security decisions, pp. 709-729. Cambridge University Press, (2017), Cambridge.

[13] J. Zhuang, V. M. Bier, and O. Alagoz. *"Modeling secrecy and deception in a multiple-period attacker–defender signaling game"*, European Journal of Operational Research, 203, pp. 409-418, (2010).

[14] V. M. Bier and N. Haphuriwat. *"Analytical method to identify the number of containers to inspect at U.S. ports to deter terrorist attacks"*, Annals of Operations Research, 187, pp. 137-158, (2011).

[15] K. Hausken and J. Zhuang. *"The timing and deterrence of terrorist attacks due to exogenous dynamics"*, Journal of the Operational Research Society, 63, pp. 726-735, (2012).

[16] V. R. R. Jose and J. Zhuang. *"Technology adoption, accumulation, and competition in multiperiod attacker-defender games"*, Military Operations Research, 18, pp. 33-47, (2013).

[17] X. Shan and J. Zhuang. *"Hybrid defensive resource allocations in the face of partially strategic attackers in a sequential defender–attacker game"*, European Journal of Operational Research, 228, pp. 262-272, (2013).

[18] N. S. Dighe, J. Zhuang, and V. M. Bier. *"Secrecy in defensive allocations as a strategy for achieving more cost-effective attacker deterrence"*, International Journal of Performability Engineering, 5, pp. 31, (2016).

[19] J. Xu and J. Zhuang. *"Modeling costly learning and counter-learning in a defender-attacker game with private defender information"*, Annals of Operations Research, 236, pp. 271-289, (2016).

[20] V. M. Payyappalli, J. Zhuang, and V. R. R. Jose. *"Deterrence and risk preferences in sequential attacker–defender games with continuous efforts"*, Risk Analysis, 37, pp. 2229-2245, (2017).

[21] P. Guan, M. He, J. Zhuang, and S. C. Hora. *"Modeling a multitarget attacker–defender game with budget constraints"*, Decision Analysis, 14, pp. 87-107, (2017).

[22] J. Zhang, J. Zhuang, and V. R. R. Jose. *"The role of risk preferences in a multi-target defender-attacker resource allocation game"*, Reliability Engineering & System Safety, 169, pp. 95-104, (2018).

[23] Z. Xu and J. Zhuang. *"A study on a sequential one-defender-N-attacker game"*, Risk Analysis, 39, pp. 1414-1432, (2019).

[24] J. Zhang and J. Zhuang. *"Modeling a multi-target attacker-defender game with multiple attack types"*, Reliability Engineering & System Safety, 185, pp. 465-475, (2019).

[25] J. Zhang, Y.Wang, and J. Zhuang. *"Modeling multi-target defender-attacker games with quantal response attack strategies"*, Reliability Engineering & System Safety, 205, (2021).

[26] P. McNamee and J. Celona, *"Decision analysis for the professional,"* 4th ed., SmartOrg, Inc, (2009), Menlo Park, CA.

[27] E. R. Johnson and S. N. Tani, *"Perform probabilistic analysis and identify insights,"* In G. S. Parnell, T. A. Bresnick, S. N. Tani, and E. R. Johnson (Eds.), Handbook of decision analysis, pp. 248-290. John Wiley & Sons, Inc., (2013), Hoboken, NJ.

[28] R. T. Clemen and T. Reilly, *"Making hard decisions with decision tools,"* Cengage Learning, (2013), Boston, MA.

[29] R. L. Dillon, W. J. Burns, and R. S. John. *"Insights for critical alarm-based warning systems from a risk analysis of commercial aviation passenger screening"*, Decision Analysis, 15, pp. 154-173, (2018).

[30] N. Scurich and R. S. John. *"Perceptions of randomized security schedules"*, Risk Analysis, 34, pp. 765-770, (2014).

[31] J. Cui, H. Rosoff, and R. S. John. *"Deterrence of cyber attackers in a three-player behavioral game,"* In S. Rass, B. An, C. Kiekintveld, F. Fang, & S. Schauer (Eds.), Decision and game theory for security: GameSec 2017, pp. 718-736. Springer, (2017), New York.

[32] S. A. Kusumastuti, J. Blythe, H. Rosoff, and R. S. John. *"Behavioral determinants of target shifting and deterrence in an analog cyber-attack game"*, Risk Analysis, 40, pp. 476-493, (2020).

[33] G. Ridinger, R. S. John, M. McBride, and N. Scurich. *"Attacker deterrence and perceived risk in a Stackelberg security game"*, Risk Analysis, 36, pp. 1666-1681, (2016).

[34] J. Cui, T. Nguyen, J. Pita, and R. S. John. *"Methods for addressing the unpredictable real-world element in security,"* In A. Abbas, M. Tambe, & D. von Winterfeldt (Eds.), Improving homeland security decisions, pp. 574-603. Cambridge University Press, (2017), Cambridge.