SCIENCE & GLOBAL SECURITY
2016, VOL. 24, NO. 1, 3–21
http://dx.doi.org/10.1080/08929882.2016.1127031

Routledge
Taylor & Francis Group

A Game Theoretic Approach to Nuclear Safeguards Selection and Optimization

Rebecca M. Ward[a] and Erich A. Schneider[b]

[a]Nuclear Forensics, The University of Texas at Austin, Austin, TX, USA; [b]Nuclear and Radiation Engineering Program, The University of Texas at Austin, Austin, TX, USA

ABSTRACT

This article presents a novel application of an inspection game to find optimally efficient nuclear safeguard strategies. It describes a methodology that allocates resources at and across nuclear fuel cycle facilities for a cost-constrained inspectorate seeking to detect a state-facilitated diversion or misuse. The methodology couples a simultaneous-play game theoretic solver with a probabilistic model for simulating state violation scenarios at a gas centrifuge enrichment plant. The simulation model features a suite of defender options based on current International Atomic Energy Agency practices and an analogous menu of attacker proliferation pathway options. The simulation informs the game theoretic solver by calculating the detection probability for a given inspector-proliferator strategy pair. To generate a scenario payoff, it weights the detection probability by the quantity and quality of material obtained. Using a modified fictitious play algorithm, the game iteratively calls the simulation model until Nash equilibrium is reached and outputs the optimal inspection and proliferation strategies. The value the attacker places on material quantity and quality is varied to generate results representative of states with different capabilities and goals. Sample model results are shown to illustrate the sensitivity of defender and attacker strategy to attacker characteristics.

ARTICLE HISTORY
Received 7 February 2014
Accepted 27 July 2015

Introduction

Concern over nuclear proliferation has elevated in concert with increased global interest in civilian nuclear power and the spread of commercial fuel cycle technologies. This confluence of factors has placed heavy demands on International Atomic Energy Agency (IAEA) Department of Safeguards, the organization tasked with verification of peaceful nuclear activities. Traditionally IAEA safeguards have been applied in a prescriptive manner according to an established set of guidelines. Safeguards implementation has been largely transparent to the states except for random on-site inspections. The regime thus places high demand on physical inspections, which are costly and manpower-intensive.

CONTACT Erich A. Schneider ✉ eschneider@mail.utexas.edu 📇 Associate Professor, Nuclear and Radiation Engineering Program, The University of Texas at Austin, 1 University Station C2200, Austin, TX 78712, USA.

Budget constraints have spurred efforts to increase IAEA efficiency through the development of tools to aid in resource allocation decision-making. Many such tools focus on diversion pathway analysis and are based on probabilistic techniques.[1] While probabilistic techniques are valuable for describing fundamentally random events, like natural disasters, their application to adversarial problems has come under scrutiny. Common criticisms include the fact that data are too scarce for many security problems to adequately characterize the threat or consequences of an attack, and the notion that probabilistic techniques may not fully capture the behavior of intentional actors like a malevolent state or terrorist.[2] Intentional actors represent a special class of threat, because they possess the ability to observe defenses and adjust their actions accordingly.[3] Cox voices this skepticism in his work, criticizing especially the use of chance nodes in fault tree analysis to model adversary decisions by arguing that these decisions are chosen based on adversary judgment, not governed by chance. Cox and others suggest that a game theoretic approach to intelligent risk analysis may be more appropriate. In light of this perceived weakness in the existing body of work, this article presents a game theoretic methodology to explore optimal IAEA resource allocation strategies for detecting illegal state behavior at a safeguarded gas centrifuge enrichment plan (GCEP).

Previous work

Examples abound of the use of probabilistic simulation techniques for diversion pathway analysis. One example is the Integrated Safeguards System Analysis Tool (LISSAT) developed at Lawrence Livermore National Laboratory. It is continuous-time model for evaluating safeguards system effectiveness at fuel cycle facilities that uses a digraph fault tree structure to examine possible points for safeguards system failure given different diversion events.[4] Another example is a Markov-model based proliferation assessment tool developed at Brookhaven National Laboratory.[5] The model features both intrinsic and extrinsic barriers to proliferation, including a suite of IAEA safeguards options. It evaluates metrics of interest, including minimum time to and cost of proliferation, the detection probability, and technical difficulty of the diversion pathway.

These and other methods examine vulnerable proliferation pathways at a single facility. Additional work characterizes proliferation resistance across multiple facilities in a fuel cycle system. A Proliferation Resistance & Physical Protection (PR & PP) evaluation methodology, developed by an expert group of the Generation IV International Forum,[6] focuses on evaluating the proliferation resistance of a nuclear energy system as a whole relative to other nuclear energy systems. The outcome is evaluated using a multi-attribute utility analysis, which includes detection probability, proliferation time, and "safeguardability."

These probabilistic tools, while useful for signature development for a diversion event or to assess relative proliferation resistance of different fuel cycle systems, may not be well suited for resource allocation decision making. They rely heavily on user input for the diversion scenarios modeling, and thus the scenarios are constrained to those imagined by the analysts. More important, these techniques fail to capture the intelligent and adaptive nature of the adversary, namely his ability to observe static or transparent defenses and change his strategy accordingly.

Game theory is a popular technique for modeling adversarial situations because of its ability to emulate rational human cognition and behavior, and as such, it has historically been applied to safeguards and inspection situations.[7] For example, Avenhaus presents a game theoretic treatment of data and material accountancy verification at nuclear facilities.[8] An analysis of attribute sampling across multiple strata is considered, and the mathematical formulation for optimal inspector strategy is given. This formulation is the foundation for the current IAEA attribute-sampling paradigm. Defensive resource allocation across multiple facilities is examined, and optimal inspector and inspectee strategies are given for a scenario with a small number of facilities.

Kilgour and Avenhaus use game theory and decision theory to examine the cost-effectiveness of IAEA inspections and recommend strategies to improve efficiency.[9] The work establishes that a state's motivation to violate depends on political parameters—namely, the penalty the state perceives for detected illegal behavior and the reward the state perceives for undetected illegal behavior—as well as a technical parameter, inspection effectiveness.

This previous work is largely theoretical in nature, presenting the mathematical formulation for strategies and detection probabilities. The objectives of the past work generally did not require that individual safeguards be depicted at a realistic level of detail. Furthermore, the complex nature of the game formulation has led others to limit its scope to allow for the calculation of Nash equilibrium.[10]

In an application of the game theoretic approach to a specific proliferation strategy, Brown et al. present a more applied two-stage, max–min Stackelberg game representing an interdictor trying to maximally delay a proliferator, who is trying to produce a first batch of fissile material.[11] The model assumes that the proliferator observes the interdictor's defense strategy and adjusts his strategy accordingly. The incorporation of a detailed project management sub-model, which is coupled with the game model for optimization, allows the diversion scenario to be described using a parametric model. The simulation tool in this work extends Brown's approach of using a sub-model to generate outcomes for each interdictor/ proliferator strategy.

An alternative approach to modeling intelligent adversary behavior is Agent-Based Modeling (ABM), which seeks to explain and predict group dynamics by modeling individual behavior and interactions. The Bayesian Agent Based Modeling (BANE) tool has been developed to model the interaction between defensive and offensive nonproliferation agents and explore the interplay between demand-side and supply-side factors that may influence a state's propensity to pursue nuclear weapons.[12] While this model nicely captures the adaptive nature of both agents, it takes a broader, network-level view, as opposed to the more detailed, facility-level approach presented here.

Game model

In this work, the game is modeled as a two-person zero-sum (TPZS) simultaneous play (Cournot) game. A TPZSG is used to model the interaction between two players with diametrically opposing goals—in this case, the attacker seeks to minimize the payoff and the defender seeks to maximize the payoff. While a TPZSG is an imperfect model for the complex interaction between a state contemplating illegal behavior and an international inspectorate, it is employed here because of its simplicity and flexibility. A TPZSG is solvable with limited mathematical and/ or computational expense, which allows for additional richness and complexity to be built into the simulation model. In addition, the use of a TPZSG represents a conservative assumption, because it provides for the defender selecting a strategy in the worst-case scenario, that is, against the attacker's most dangerous strategy.

In a simultaneous play game, both players have full knowledge of the strategy options available to the other player, but each player must commit to his strategy before observing to what strategy the other player commits. The assumption of perfect knowledge represents a modeling idealization and simplification; in reality, it is unlikely (and undesirable) that an adversary would have perfect knowledge of the options available to the defender and his chance of defeating each defender strategy. Yet this assumption is more realistic for insider adversaries, like a proliferant state, than for outsiders, given insiders' knowledge of security measures and operational procedures. As with the TPZSG, this assumption is generally a conservative one, as it finds the optimal defender strategies against a more informed and thus more capable adversary.

The game is solved using the fictitious play (FP) algorithm. Fictitious play is a myopic learning algorithm first introduced by Brown for finding the value of a TPZSG.[13] Fictitious play is an alternative to the standard Simplex method and can be advantageous for large linear systems.[14] FP was employed in this work because unlike the Simplex algorithm, it eliminates the requirement to pre-populate the payoff matrix, which dramatically reduces the number of simulation calls needed to solve the game. In the FP process, each player assumes her opponent is playing a stationary strategy, and the two players engage in an iterative finite game. In each round a player chooses her myopic best response to the distribution of strategies played by her opponent up to that point; that is, she selects the response that will maximize her expected payoff in the next round of play. Julia Robinson showed that all TPZSGs converge to the Nash equilibrium value as the number of iterations approach infinity.[15]

Figure 1 depicts a flowchart of the game model and its interaction with the simulator. Defender and attacker strategies are indexed by i and j, respectively, where $i \in [0, I]$ and $j \in [0, J]$. Pure attacker and defender strategies[16] are denoted by y_j and x_i, respectively. The payoff for defender strategy x_i and attacker strategy y_j is v_{ij}. $\mathbf{x}$ is an I-element vector that holds the defender's mixed strategy history; the *ith* element of $\mathbf{x}$ is incremented when the defender plays pure strategy x_i, and the values in the vector are re-normalized such that the I elements in $\mathbf{x}$ sum to 1. $\mathbf{y}$ is the analogous attacker mixed strategy history.

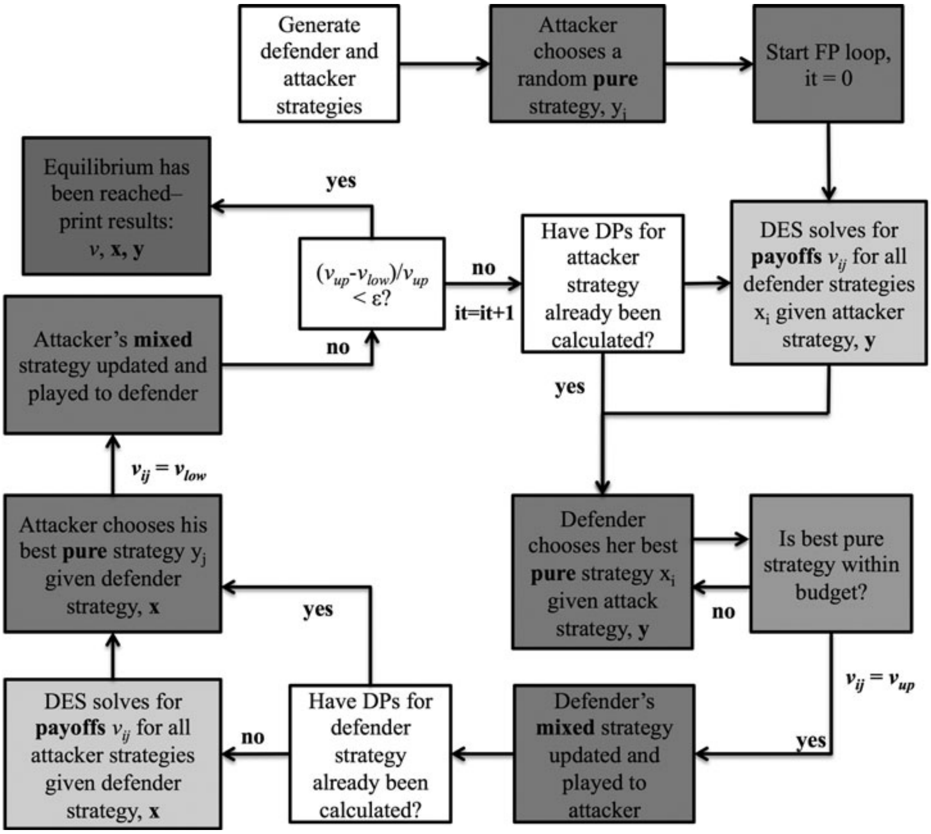
Generate defender and attacker strategies
Attacker chooses a random **pure** strategy, y_j
Start FP loop, it = 0
DES solves for **payoffs** v_{ij} for all defender strategies x_i given attacker strategy, **y**
Have DPs for attacker strategy already been calculated?
yes
no
it=it+1
$(v_{up}-v_{low})/v_{up} < \varepsilon$?
yes
no
Equilibrium has been reached–print results: v, **x**, **y**
Attacker's **mixed** strategy updated and played to defender
$v_{ij} = v_{low}$
Attacker chooses his best **pure** strategy y_j given defender strategy, **x**
yes
Defender chooses her best **pure** strategy x_i given attack strategy, **y**
Is best pure strategy within budget?
no
$v_{ij} = v_{up}$
yes
Defender's **mixed** strategy updated and played to attacker
Have DPs for defender strategy already been calculated?
no
DES solves for **payoffs** v_{ij} for all attacker strategies given defender strategy, **x**

Figure 1. Flowchart of model logic.

The FP algorithm is initiated by the attacker randomly choosing and playing pure strategy, $y_j^{(0)}$ ("Start FP loop" box in Figure 2). The simulator is called and calculates the payoffs v_{ij} for all defender strategies x_i, given the attacker's strategy $\mathbf{y}$, and these values are stored in the payoff matrix. Knowing the payoffs for all defender strategies that can be played in response to $y_j^{(0)}$, the defender then chooses the pure strategy response, $x_i^{(1)}$, that will maximize her payoff in the next round. After selecting the best response, the cost of the strategy is checked to see if the strategy is under budget. If so, the strategy is played and the defender's mixed strategy is updated. If not, the defender picks her next best pure strategy response. The defender continues to pick her next best pure strategy response until she chooses one that she can afford. Once she selects and plays her best strategy response $x_i^{(1)}$, a variable v_{up} is initialized and set equal to the value of $v_{ij}^{(1,0)}$. The simulator is then called again and calculates payoffs v_{ij} for all attacker strategies y_j in response to the defender's best pure strategy $x_i^{(1)}$. The attacker chooses his best pure strategy response, $y_j^{(1)}$, given the defender's current strategy history, **x.** The variable v_{low} is set equal to the payoff value $v_{ij}^{(1,1)}$. The attacker plays his pure strategy best response $y_j^{(1)}$ and his mixed strategy is updated accordingly. This constitutes one fictitious play loop, and convergence is checked. The model is said to have converged when the convergence criterion $(v_{up}\text{-}v_{low})/v_{up} < \varepsilon$ is met. For the results presented in this article, $\varepsilon = 0.001$. At convergence, the

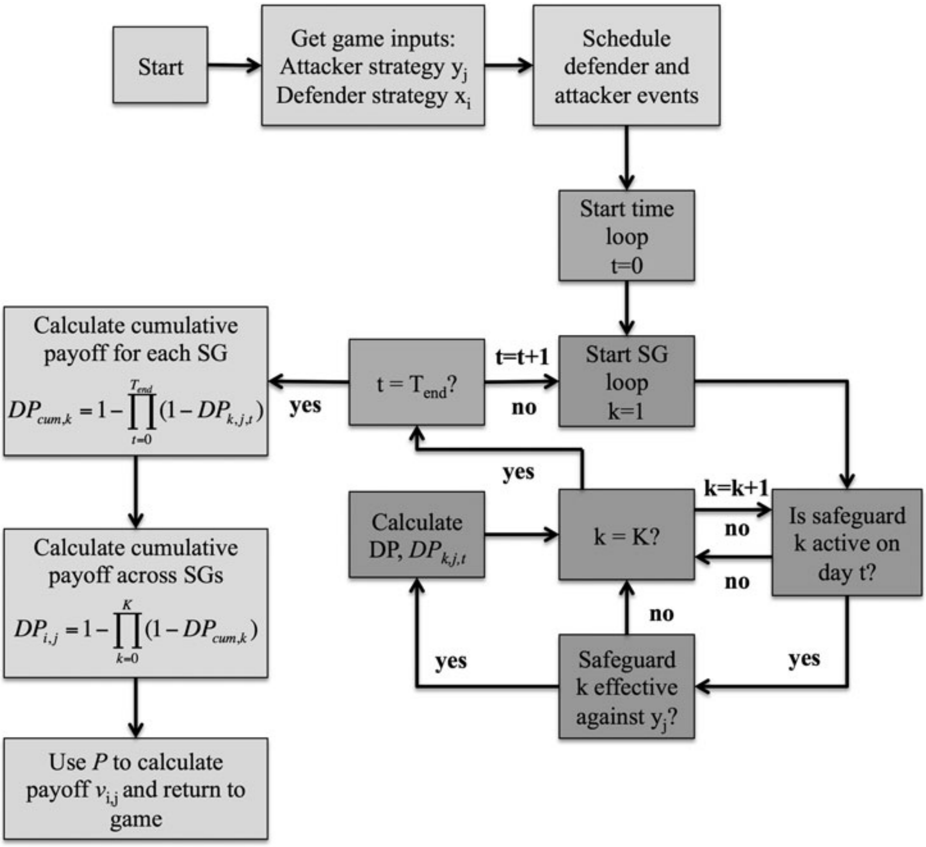
Start
Get game inputs: Attacker strategy y_j Defender strategy x_i
Schedule defender and attacker events
Start time loop t=0
Start SG loop k=1
t=t+1
$t = T_{end}$?
no
yes
Calculate cumulative payoff for each SG
$DP_{cum,k} = 1 - \prod_{t=0}^{T_{end}} (1 - DP_{k,j,t})$
Calculate cumulative payoff across SGs
$DP_{i,j} = 1 - \prod_{k=0}^{K} (1 - DP_{cum,k})$
Use P to calculate payoff $v_{i,j}$ and return to game
yes
k=k+1
Is safeguard k active on day t?
no
no
k = K?
Calculate DP, $DP_{k,j,t}$
no
yes
yes
Safeguard k effective against y_j?

Figure 2. Simulation logic.

mixed strategies **x** and **y** are the equilibrium defender and attacker strategies, respectively, and $v_{up} = v_{low} = v$, the equilibrium value of the game. If convergence is not yet achieved, control is returned to the FP loop.

Simulation model logic

Figure 2 details the logic flow for the simulation model. When calling the simulation, the game passes all defender and attacker strategy information needed to define a strategy pair or scenario. The simulation uses these inputs to create schedules of defender and attacker events for the course of the simulation period. The length of the simulation period is determined by the attacker strategy. Day 1 of the simulation is the defined as the day the attacker begins his malevolence. The attacker chooses the duration of the attack (unless the attack is a single event, in which case it proceeds for only one day). The simulation period extends after the end of the attack to provide the defender time to detect missing material and place the facility in an "alert state."[17] For the results presented here, the post-diversion detection time was set to thirty days to correspond to the IAEA timeliness goals for the detection of a significant quantity of unirradiated highly enriched uranium (HEU).[18]

Table 1. Annual Throughput for 465,000 Kg-SWU GCEP[1].

Material	Mass UF_6 (kgU)	Annual cylinder throughput	Number of cylinders assumed in storage at any given time
Feed	552 500	65	13
Product	63 390	41	3
Tails	489 200	57	—

[1]Calculated directly; calculations validated against enrichment calculator at uxc.com.

A simulation day begins with a loop through all deployed safeguards, k, to ascertain whether each is active on day $t = 1$. Here$k \in [0, K]$ indexes across all safeguards. Once each active safeguards measure is identified, a check is conducted to see if that measure is effective against the active attacker strategy. If the safeguards measure k is effective against attacker strategy y_j, the daily detection probability for the pair, $DP_{k,j,t}$, is calculated. The algorithms used to calculate DP depend on the safeguard-attacker option pair; they are summarized in the GCEP Simulation Model section of this article, and detailed descriptions are available in Appendix B and reference.[19] After the payoffs for all active safeguards on a given day have been calculated, this process is repeated every subsequent day until the simulation time has been exhausted. The cumulative payoff for each safeguards measure over the simulation period, $DP_{cum,k}$, is calculated by taking the multiplicative sum across all the days $t \in [0, T_{end}]$, as shown in Figure 2. An overall scenario DP, $DP_{i,j}$, is calculated by combining the individual safeguards measure DPs, and the scenario payoff, $v_{i,j}$, is calculated using a payoff function, P, to weight the scenario DP by value of the material in the scenario. Finally, the scenario payoff value itself is returned to the game.

A gas-centrifuge enrichment plant (GCEP) with an annual capacity of 465,000 kg-SWU is modeled. A GCEP was selected for this case study because of its perceived proliferation risk[20] and the resources the IAEA has historically dedicated to safeguarding this type of facility.[21] The facility was chosen to represent a plant that might be present in a nation with a burgeoning indigenous fuel cycle. The plant uses natural uranium feed with uranium-235 enrichment of 0.711% and enriches product to 4.5% uranium-235, producing tails with 0.22% enrichment. Annual material throughput under normal operating conditions is shown in Table 1. The third column of Table 1 gives the number of each type of cylinder assumed to be in storage at any given time. It is assumed that approximately 84 days' worth of feed is kept on site at all times and that the operator is required to hold all product cylinders for a period of 28 days (at least one inspection cycle).[22]

The defender and attacker options used in the GCEP simulation model are displayed in Table 2 and summarized below. Shaded cells indicate that the safeguards measure is effective against the corresponding attacker option. The attacker options and defender options and algorithms used to calculate the detection probability for each defender-attacker strategy pair are described in detail in Appendix A and B,

Table 2. Defender-Attacker Strategy Pair Summary Table for Enrichment Facility. X Indicates Defender Option is Effective Against Attacker Option.

	Attacker Options					
Defender Options	1	2	3	4	5	6
A	X	X	X			X
B		X				
C				X	X	
D				X	X	
E	X	X				
F			X	X		
G				X	X	
H			X	X	X	X
I				X	X	

respectively. It should be noted that the facility and inputs used to calculate detection probabilities are notional, albeit representative. An effort was made to accurately capture the relative effect of different defender and attacker parameters on DP (i.e., stealing larger quantities of material is more likely to be detected than stealing smaller quantities of material); however, the values are intended to allow for comparison of options and are meaningful only in a relative sense. When all possible permutations of the different options and sub-options available to the players are enumerated, the GCEP simulation model defines a total of 246,645 defender options and 321 attacker options.

The defender and attacker strategies presented in this article are single proliferation actions at a single facility; however, the strength of the game theoretic method is that it is extensible to system-level modeling and analysis. Though not presented here, this model has demonstrated capability modeling multi-facility systems, allowing the attacker to choose his point of attack and consequently forcing the defender to allocate resources across multiple facilities. As such, it could be used to model complete state-level diversion or acquisition pathways, including threats from undeclared facilities. The costs and detection methods for undeclared facilities differ from those of declared facilities, but they could in principle be incorporated into the model by simply integrating the appropriate simulation inputs. Listner and Canty offer four possibilities for modeling detection probabilities at undeclared facilities, including a Bayes and frequentist approach that could be incorporated into the simulation model.[23]

Attacker options

The attacker has six attack categories from which to choose, representing three major types of attack: diversion of declared product, misuse of the facility to enrich above declared levels, and production of undeclared product from undeclared feed. An attacker strategy is comprised of only one attacker option and its defined parameters, which the attacker selects from a set of discrete options.

Table 3. Attacker Options and Associated Parameters.

Attacker option	Tunable parameters [allowable values]	Description
1. Diversion of cylinder (cyltheft)	Number of cylinders [1, 2, 3] Area [feed, storage]	Attacker diverts cylinder(s) from feed or product storage in one-time attack.
2. Diversion of some material from cylinder (matcyl)	Frequency [1, 7, 30 days-1] Durations [7, 30, 360 days] Total mass [40, 110, 775 kg] Number of cylinders [1, 2, 3]	Attacker diverts material from cylinder(s) in product storage in a continuous attack.
3. Diversion of material from cascade (matcasc)	Frequency [1, 7, 30 days-1] Durations [7, 30, 360 days] Mass removed per cascade [0.010, 0.100 kg] Number of cascades [1, 6, 30]	Attacker diverts some material at product enrichment from cascade(s) in a continuous attack.
4. Re-piping cascade (repiping)	Durations [7, 30, 360 days] Fraction cascades dedicated [0.0167, 0.1, 0.5] Product enrichment [0.197, 0.50, 0.90]	Attacker re-pipes the cascades in a one-time attack and then continues to produce material with the illegal cascade configuration daily to produce material enriched above declared values.
5. Recycling material through cascade (recycle)	Frequency [1, 7, 30 days-1] Durations [7, 30, 360 days] Number of cascades [1, 6, 30] Product enrichment [0.197, 0.50, 0.90]	Attacker recycles material through the cascade to produce material enriched above declared values.
6. Undeclared feed (udfeed)	Frequency [1, 7, 30 days-1] Durations [7, 30, 360 days] Number of cascades [1, 6, 30]	Attacker feeds undeclared material through the cascade to produce undeclared feed at product enrichment.

Table 3 lists each attacker option, the relevant parameters, and a brief description. The parenthetical label after the attack option is the label used to refer to the option in model results.

Defender options

Table 4 enumerates each defender option and relevant tunable parameters. As with the attacker options, the parenthetical label appearing after the name of the defender option gives the label used to refer to the option in model output; these labels are used in the Results section. A defender strategy is comprised of any number of active safeguards measures, depending on how many safeguarding options the defender elects to purchase. Detailed descriptions of each safeguards measure are given in Appendix B.

Exogenous detection probabilities

In addition to the safeguards listed above, cost-free, exogenous sources of detection capability are incorporated into the model using a background DP. Background DP serves as a proxy for all other safeguards measures and sources of detection not explicitly considered, including the increased detection capability that intelligence and open source information offer, generally *at no cost to the inspector.* The background DP is a daily probability and is attacker strategy-specific. This implementation is intended to represent the reality that intelligence and open source information is better suited to detect certain diversion/misuse scenarios.

Table 4. Defender Options and Tunable Parameters.

Defender option	Tunable parameters [allowable values]	Description
A. Inspection (inspection)*	Frequency [7, 28 days-1] Team size [small, large] False Alarm Probability (FAP) [0.01, 0.001]	A basic inspection is comprised of physical inventory of cylinders in storage, mass balance verification, and review of logged video surveillance images.
B. Passive seal verification (pseals)†	Frequency [7, 28 days-1] Fraction seals verified [0.5, 1]	Checking integrity of passive seals to determine if tampering occurred.
C. Non-Destructive Assay (nda)†	Frequency [7, 28 days-1] FAP [0.01, 0.001]	Gamma spectroscopy to determine enrichment of material.
D. Destructive Assay (da)†	Frequency [7, 28 days-1] Number of samples [1, 3]	Take samples and send to lab for highly accurate but time consuming analysis to determine isotopics.
E. Review of transmitted video images	Team size [small, large]	Video is remotely transmitted and automatically reviewed for anomaly detection.
F. Active seal verification (aseals)	Fraction cascades sealed [0.5, 1]	Application of active seals to automatically alert if tampering occurs.
G. Continuous Enrichment Monitoring (cemo)	FAP [0.01, 0.001] Count time [300, 3600 s]	Continuous, online, go-no go enrichment monitoring to detect material in cascade with enrichment above 20%.
H. Visual Inspection (DIV)††	Frequency [7, 28 days-1]	Visual inspection of cascade hall for anomalies (i.e., suspiciously placed cylinders or re-piping).
I. Environmental Sampling (ES)††	Frequency [7, 28 days-1] Number of samples [6, 12]	Swipes taken in cascade hall and sent to lab for destructive analysis to provide isotopic information.

*Basic inspection; †"Add-on"—can be added to basic inspection up to as frequently as basic inspection occurs; ††Cascade hall inspection—grants inspector access to cascade hall.

Safeguards costs

A method was formulated to allocate relative costs to each safeguards measure. These costs are estimates based on available information about the necessary technology or manpower needs. The cost values used in the model, referred to as "simulation dollars" (s$), are the based upon the estimated real values of selected safeguards divided by 100. For example, a piece of equipment that costs $1000 costs 10 simulation dollars. This paradigm is used for convenience and to emphasize that the costs here retain meaning in a relative sense, but are not claimed to be faithful to the actual absolute costs.

The cost associated with each safeguards measure has two components: capital and operations and maintenance (O&M). Capital costs are amortized over the serviceable lifetime of the equipment. These are one-time costs incurred for large pieces of equipment, such as a mass spectrometer. O&M costs fall into two categories: fixed and variable. Fixed O&M costs are associated with the upkeep of the equipment and are incurred whether the equipment is used regularly or not. Variable O&M costs are costs that the defender pays when he uses the service, such as analyzing a sample, assessing surveillance feed, or inspecting a facility. The per-item cost of certain safeguards is also considered a variable O&M cost, such as the cost of a seal. The total cost for a safeguards measure is the sum of the annual equipment, fixed, and variable O&M costs.

Table 5. Enrichment Safeguards Cost Summary.

Safeguards measure	Capital cost (s$/year)	Fixed O&M (s$/year)	Variable O&M (s$/year)		Total fixed cost (s$/year)
			Manpower	Other	
Insp- Inventory	0	0	10/insp	0	0
Insp- Mass balance	12	0.24	0	0	12.24
Insp- Video logged	16.50	1.65	0	0	18.15
Passive seals	0	0	2/insp 2.50/batch	0.01/seal	0
NDA	3.6	0.07	2/insp	0	3.67
DA	8.93	0.18	2/insp 2.50/batch	0	9.11
Video transmitted	32.50	3.25	0.60/day	0	35.75
Active seals	0	0	0.60/day	0.50/seal	0
CEMO	18	1.80	0.60/day	0	19.80
Visual inspection	0	0	30/insp	0	0
ES	8.93	0.18	6/insp 5/batch	0	9.11

Table 5 shows the costs associated with each safeguards measure. A detailed description of the assumptions used to arrive at these values is given in Appendix C. The total costs for all of the possible defender strategies in the model range from 0-5900 s$.

Payoffs

The payoff to the defender and attacker for a given strategy pair is the detection probability weighted by the quantity and attractiveness of the material obtained. The material attractiveness is valued using Bathke's Figure of Merit (FOM) method[24] for an advanced proliferant state or a sub-national group unconcerned with yield. The FOMs for all material available to the attacker in the enrichment simulation are given below in Table 6. For a detailed description of how the FOM values were calculated, please see the reference from note 18. A FOM value for natural and 4.5% enriched uranium could not be calculated because these materials have an infinite bare sphere critical mass; however, values were assigned to characterize the attractiveness of these materials relative to the other enriched uranium products.

The FOM value is combined with the material quantity Q [kg] and DP using the payoff function, P, given in Equation 1. Here α is a weighting factor that describes

Table 6. FOM Values for Enrichment Facility.

Enrichment	FOM
0.711%	0.033[1]
4.5%	0.1
19.7%	0.991
50%	1.69
90%	2.15

[1] A FOM-like value of 0.1 was assigned to 4.5% enriched material, and the value of 0.033 for natural uranium was assigned one-third of that value based on the fact that enriching 1 kg of natural uranium to 90% requires approximately three times the enrichment capacity as enriching 4.5% enriched material to 90%.

the degree to which the attacker is motivated by high-value material, or the "material premium" parameter. The e parameter ensures that the payoff becomes very large, but not infinite, for DP = 1 scenarios and is assigned the value of 0.001. The payoff function becomes asymptotically large as the DP approaches unity, which is undesirable to the attacker, who seeks to minimize the payoff. Thus despite the attacker's incentive to obtain high-value material, he will reject any strategy that results in certain detection, if alternatives are available. The payoff takes the maximum value of 1.0 in a breakout scenario, meaning two conditions are met: (1) the attacker obtains the best possible material, and (2) the scenario DP is one. The notation $\left(FOM_{y\prime} \cdot Q_{y\prime}\right)$ indicates the maximum possible material utility for the attacker. The payoff is normalized by the maximum possible material value available to the attacker to eliminate an artificial drop in payoff value as alpha increases.

$$P = \frac{DP}{(FOM \cdot Q)^{\alpha}} \cdot \frac{1}{(1 + e - DP)} \cdot e \cdot (FOM_{y'} \cdot Q_{y'})^{\alpha} \qquad (1)$$

Sample model results

Three illustrative results are presented here to demonstrate some of the model's capabilities: defender and attacker strategy sensitivity to attacker characteristics, the efficient frontier, and defender strategy sensitivity to exogenous detection. The first result shows the sensitivity of defender and attacker strategy to the premium the attacker places on material value. In order to perform this analysis, the alpha value shown in Equation 1 was varied from 0 to 0.8, and the equilibrium strategies were calculated. Note that $\alpha = 0$ is not a realistic scenario, because it ascribes equal utility to all materials and amounts from natural uranium to HEU; however, $\alpha = 0$ serves as the limiting case of an extremely conservative attacker. Alpha values were varied only up to 0.8 because at this value the attacker has already committed to a single attack strategy that is dominated by his desire for high-value material.

Figure 3 and Figure 4 show attacker and defender strategy, respectively, as a function of alpha for a budget level of B = 200. The numbers assigned to each strategy are

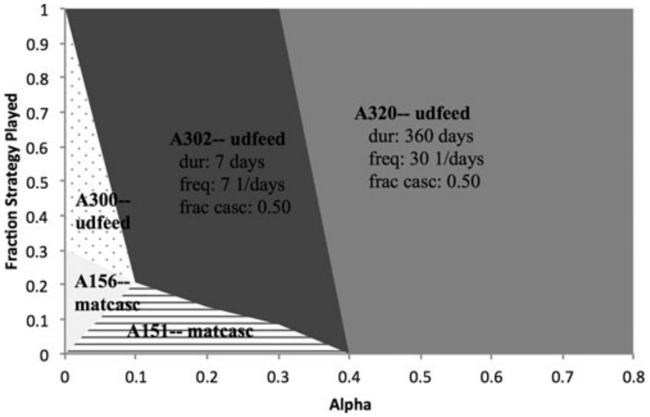

Figure 3. Attacker strategy as a function of alpha (B = 200).

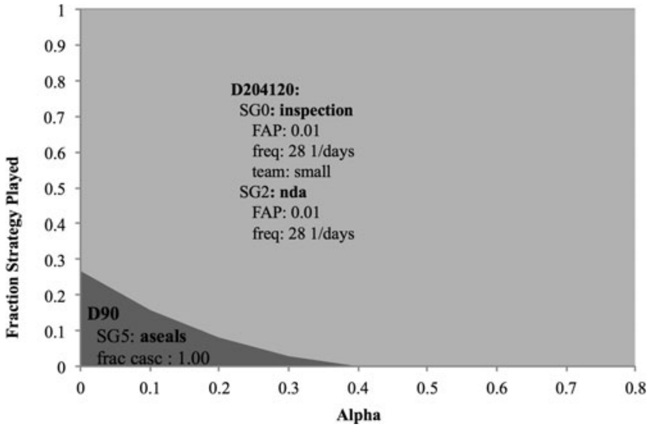

Figure 4. Defender strategy as a function of alpha (B = 200).

unique identifiers of the defender strategies and attacker strategies. The title next to the number indicates the type of safeguards measure deployed or the type of attack being perpetrated. The fraction of each pure strategy played is given on the vertical axis. A pure strategy is comprised of only one attacker or defender option, while a mixed strategy contains membership from more than one option. It represents a randomization between the pure strategies of which it consists.

In the low alpha region, both the defender and attacker play mixed equilibrium strategies. Here material attractiveness and quantity do not strongly affect the payoff, so the attacker chooses a relatively low-risk strategy that is difficult to detect but yields little material value. Specifically the attacker plays a mixed strategy of producing undeclared product from undeclared feed and stealing low-enriched uranium directly from the cascade. Conceptually the mixed strategy represent a randomization of strategy options; for example, at alpha = 0 there is a 70% chance that the attacker will produce undeclared product and 30% chance that he will steal material directly from the cascade. Though it is not immediately evident from the figure, it can be seen from the strategy descriptions in Table 7 that the strategies played by the attacker become increasingly brazen as alpha increases and the attacker becomes more incentivized by material value and quantity. For example, A302 and A300 are both weeklong attacks wherein the attacker produces undeclared feed, but for strategy A300, the attacker produces undeclared feed only once during the period, and for strategy A302 the attacker produces undeclared feed daily. At alpha = 0.4 the attacker is sufficiently motivated by desire for large quantities of material that he

Table 7. Attacker Strategy Descriptions.

Strategy	Parameter 1	Parameter 2	Parameter 3	Parameter 4
A300- udfeed	dur = 7 days	freq = 7 days^{-1}	fraction = 0.0167	
A156- matcasc	dur = 7 days	freq = 7 days^{-1}	fraction = 0.0167	mass = 0.010 g
A302- udfeed	dur = 7 days	freq = 1 days^{-1}	fraction = 0.50	
A151- matcasc	dur = 7 days	freq = 1 days^{-1}	fraction = 0.0167	mass = 0.100 g
A276- recycle	dur = 360 days	freq = 1 days^{-1}	fraction = 0.50	x_p = 0.197
A320- udfeed	dur = 360 days	freq = 30 days^{-1}	fraction = 0.50	

switches to a more aggressive pure strategy of undeclared feed production while retaining a non-zero evasion probability.

Figure 4 shows that the defender also plays a mixed strategy at low alpha values and then switches to a pure strategy. At alpha = 0 and alpha = 0.1, the defender plays a mixed strategy featuring both active seals (D90), and an inspection with NDA (D204120). The mixed strategy again represents randomization: the defender plays D90 27% of the time, meaning about a quarter of the active seals applied at the facility are real seals that can relay information to a person waiting to assess an alarm, while the other three-quarters are dummy seals. The attacker can see the seals on the cascades, but cannot discriminate between real and dummy seals. The defender also plays the inspection + NDA strategy 73% of the time, meaning the defender randomly conducts only 73% of her permissible inspections. The defender uses active seals to counter material theft from the cascades, and purchases an inspection to detect undeclared feed. At high alpha values, on the other hand, the attacker places greater value on high-quality material and commits to a pure strategy of undeclared feed production. Then the defender commits to a pure inspection-only strategy, as purchasing active seals would no longer provide any detection capability.

The alpha sensitivity study is useful for understanding the results of the model in the context of real proliferation situations. In reality an adversary's preferences and capabilities (as expressed by his utility function) are difficult to know with any certainty, so the alpha sensitivity results can be used to manage this uncertainty. The alpha sensitivity study could be used in a variety of ways to inform decision-making: for example, a distribution could be placed on α to reflect prior beliefs about the attacker's utility function or capabilities. The model could then be run stochastically by performing Monte Carlo sampling from the distribution and solving for Nash equilibrium to develop a portfolio of possible defensive investments.

Fixing the value of α at 0.25, Figure 5 illustrates an "efficient frontier," the payoff as a function of budget. The step increases in payoff at budgets of 500, 1550, and 2000 s$ correspond to changes in defender strategy. These budgets elevate the

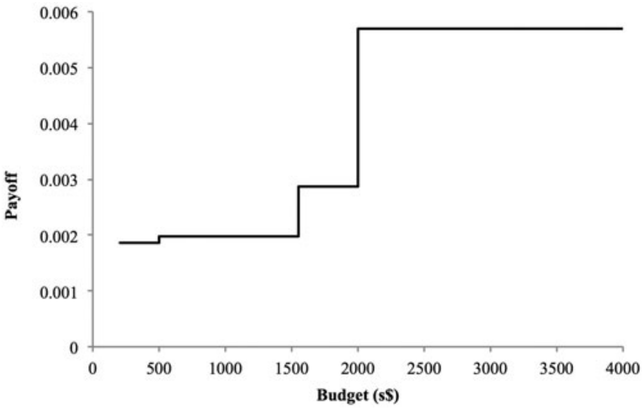

Figure 5. Payoff as a function of budget.

defender above a cost threshold and allow her to purchase some advantageous symbiotic safeguards combination that will result in increased detection capability. For example, at 1500 s$ the defender is playing a pure strategy that includes monthly inspections with a large team, NDA, active seals, and monthly cascade hall inspections (D226086). Once her budget increases to 1550 s$, the inspector is able to afford a mixed strategy that contains a randomization between all of the elements listed above and weekly (rather than monthly) cascade hall inspections (strategy: 94% D1 [weekly cascade hall inspections]; 6% D226086). The mixed strategy means that the defender does not opt to purchase a cascade hall inspection every week, but only on 94% of weeks.[25] Inspecting at this frequency is sufficient for her to deter the attacker from a more damaging strategy that could be carried out if there were no inspection. Because the attacker is playing a mixed strategy comprised largely of undeclared product production, the shift of the defender's strategy to playing primarily weekly cascade hall inspections increases the inspector's payoff, as seen in the plot, while still deterring the attacker from shifting to a different target.

This plot is important because it can serve as a guide for rational decision-making by providing information about when additional investment provides diminishing returns, as is the case in the plateau regions where additional investment does not result in increased payoff. The efficient frontier also illustrates when additional investment should be expected to increase payoff, and at what level additional resources need to be invested in order to affect the payoff .

As mentioned above, a daily background detection probability can be applied in the simulation model to serve as a proxy for exogenous detection means not explicitly modeled, including intelligence and open source information, which offers additional detection to the defender. If provided by a third party, these data streams can be cost-free or virtually cost-free to the defender. Intelligence has a non-uniform probability of detecting different types of attacks; thus, the background DP is applied non-uniformly across the attacker options. To test the sensitivity of the equilibrium strategies to the background DP, the background DP against undeclared production was systematically varied from 0.001 to 0.1, and changes in defender strategy were observed. The background DP for all other strategies was held at zero and alpha was set to 0.25. This trial was designed to mimic the real-world situation where intelligence collection may be able to detect unusual cylinder traffic into and out of a facility, as would be necessary for producing undeclared product from undeclared feed, even with no knowledge of operations inside the facility. Figure 6 shows the changes in defender strategy as a function of background DP for B = 200.

Even for a daily background DP as low as 0.1%, the defender changes the fraction of the pure strategies played in the equilibrium mixed strategy. With a daily background DP of 1%, the defender introduces a small fraction of a new pure strategy to her equilibrium mixed strategy. This change in defender strategy occurs in response to anticipated changes in attacker strategy. With the introduction of background DP, the attacker begins to shift away from undeclared production, because this is the only attacker option to which the background DP applies. When the daily background DP is as high as 5%, the attacker ceases undeclared production entirely. This

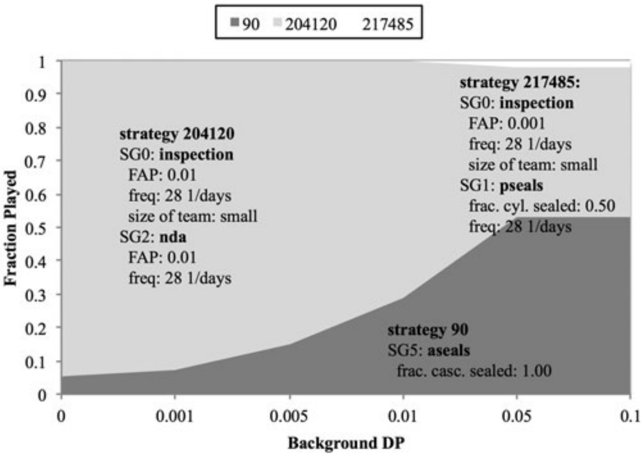

Figure 6. Defender strategy as a function of background DP for B = 200.

result has big implications for the selection of inspection strategies at low budgets; namely that the optimally efficient inspection strategy in the absence of intelligence information is not necessarily the optimally efficient inspection strategy if intelligence or open source information is available. Thus, a cost-constrained inspector must consider available reliable exogenous sources of detection in order to employ an optimally efficient strategy.

Conclusions and future work

This article presents a novel methodology and its computational implementation for using game theory to allocate safeguards activities at nuclear fuel cycle facilities. The methodology couples a game theoretic solver with a probabilistic simulation model of misuse or diversion scenarios at a GCEP. The game calls the simulation model to generate payoff values for given safeguards and attack strategy pairs, and the simulation model calculates the payoffs by weighting the detection probability for the pair by the quantity and quality of material obtained. These payoff values are returned to the game and used to populate the payoff matrix. The game is solved using a fictitious play algorithm, and the model outputs the equilibrium defender and attacker strategies as well as the equilibrium value. The methodology has been developed to allow users to input facility-specific assumptions and detection probability algorithms to generate realistic results.

The case study results show that both optimal proliferation and safeguards strategies are dependent on the attacker's own valuation of material he could potentially obtain, and by extension state capabilities, thus lending support to the necessity for state-specific nonproliferation analyses to drive safeguarding strategy decisions. Furthermore, though not presented in this article, the model can also optimize resource allocation across multiple facilities, illustrating, for instance, the percentage of her total budget the defender should invest in safeguarding multiple facilities. These two model capabilities make it a useful tool for guiding and supporting the IAEA's implementation of the State-Level Concept and information-driven safeguards.[26] This tool can provide a systematic basis for allocating safeguarding resources across multiple facilities for a state with particular characteristics.

One compelling result generated by the model is the so-called "efficient frontier," or the visual representation of scenario payoff as a function of budget. The efficient frontier traces the optimally efficient strategy at any budget, and serves as a guide to resource investment decision making by conveying information about the defender's return on investment for safeguards strategy decisions. While the results conform to the intuitive notion that increasing the defender's resource investment level generally increases the defender's payoff, they also indicate that there are certain conditions under which additional defender investment can be wasteful. One such condition is if the attacker is willing to breakout; another arises if the defender invests further resources in areas that the attacker is already deterred from attacking.

The tool can also model the sensitivity of defender strategies to exogenous sources of detection probability that are cost-free to the defender, like intelligence, to inform optimal safeguarding strategies in the presence of such information.

A significant potential application for the model developed in this article is for marginal cost analysis, particularly in the area of safeguards investment decision-making. This model could be used to perform cost sensitivity analysis for a new type of safeguards tool or technique, by determining cost above which the defender no longer selects it because the detection probability to cost ratio is too low. In a similar vein, the model could also provide an estimate for the "value" of intelligence or open source information in a specific threat environment.

Finally, one of the more pressing policy questions surrounding states that may or may not have proliferant aims is whether the state can be deterred from proliferating, and if so, at what cost. A legal behavior option could easily be implemented in this model to draw a quantitative relationship between attacker characteristics and "deterrence budget," or the investment level required by the defender to compel a state into compliant behavior or an open breach. Such an analysis would provide policy makers unique insight into how safeguards investments do or do not affect the decision made by a state to pursue an illicit weapons program.

Funding

This material is based upon work supported by the U.S. Department of Homeland Security under Grant Award Number, 2012-DN-130-NF0001. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Department of Homeland Security. Funding was also received from the Stanton Foundation and the Harvard Kennedy School's Belfer Center for Science and International Affairs.

Notes and References

1. H.A. Elayat, H.E. Lambert and W.J. O'Connell, "Systems Analysis of Safeguards Effectiveness in a Uranium Conversion Facility" (paper presented at the 45th Annual Meeting of the Institute of Nuclear Materials Management, Orlando, FL, 18–22 July 2004); H. Lambert, H. Elayat, W.J. O'Connell, L. Szytel, and M. Dreicer, "LISSAT Analysis of a Generic Centrifuge Enrichment Plant," (paper presented at the 48th Annual Meeting of the Institute of Nuclear Materials Management, Tucson, AZ, (8–12, July 2007); and M. Yue, L. Cheng, and R. Bari, "A Markov Model Approach to Proliferation-Resistance Assessment of Nuclear Energy Systems," *Nuclear Technology,* 162 (2008): 26–44.
2. The National Academies. National Research Council. Review of the Department of Homeland Security's Approach to Risk Analysis (Washington: The National Academies Press, 2010).
3. Louis Anthony Cox, Jr., "Game Theory and Risk Analysis," *Risk Analysis,* 29, 8 (2009): 1062–68, doi:10.1111/j.1539-6924.2009.01247.x.
4. Elayat et al., "Systems Analysis of Safeguards Effectiveness"; and H.A. Elayat, W.J. O'Connell, and B.D. Boyer, "Gas Centrifuge Enrichment Plant Safeguards System Modeling" (paper presented at the 47th Annual Meeting of the Institute of Nuclear Materials Management, Nashville, TN, 16–20 July 2006).
5. Yue et al., "Markov Model Approach."
6. T. Bjornard et al., "Evaluation Methodology for Proliferation Resistance and Physical Protection of Generation IV Nuclear Energy Systems: an Overview" (paper presented at the International Conferene on Probabilistic Safety Assessment and Management,New Orleans, LA, 14—18 May 2006).
7. J. Watson, Strategy: An Introduction to Game Theory, 2nd ed., (New York: W.W. Norton & Company, 2008).
8. R. Avenhaus and M.J. Canty, Compliance Quantified: An Introduction to Data Verification, (New York: Cambridge University Press, 1996).
9. Department of Foreign Affairs and International Trade Non-Proliferation, Arms Control and Disarmament Division Canada, DM Kilgour and Rudolf Avenhaus, "The Optimal Distribution of IAEA Inspection Effort: Final Report," Ottawa, 1994.
10. The Nash equilibrium of a two-person, adversarial game is the condition where neither player can become better off by changing they strategy, given that they are aware of the strategies of their adversary.
11. G.G. Brown et al., "Interdicting a Nuclear-Weapons Project," *Operations Research,* 57, 4 (2009): 866–77.
12. R. Elmore and W. Charlton, "Nuclear Nonproliferation Analysis Using Agent Based Modeling in an Entropy Empowered Intelligent Agent Bayesian Framework" (paper presented at 2014 Winter Simulation Conference, Savannah, GA, 7–10 Dec 2014; R. Elmore and W. Charlton, "Dynamic Agent Based Modeling Using Bayesian Framework for Addressing Intelligence Adaptive Nuclear Nonproliferation Analysis" (PhD diss., Texas A&M University, 2014).
13. George W. Brown, "Some Notes on Computation of Game Solutions," Rand Corporation (1949).
14. A Washburn, "A New Kind of Fictitious Play," *Naval Research Logistics* 48 (2001): 270–80.
15. Julia Robinson, "An Iterative Method of Solving a Game," *The Annals of Mathematics,* 54, 2 (1951), 296–301.
16. A pure strategy is an option that a player will always play, in contrast to a mixed strategy, where the player selects a strategy from a distribution.
17. F.A. Duran, "Probabilistic Basis and Assessment Methodology for Effectiveness of Protecting Nuclear Materials" (Ph.D. diss., The University of Texas at Austin, 2010).

18. BD Boyer, "Safeguards Approaches for Gas Centrifuge Enrichment Plants," Los Alamos National Laboratory, LA-UR-08-03736.
19. R M Ward, "A Game Theoretic Approach to Nuclear Safeguards Selection and Optimization," (Ph.D. diss., The University of Texas at Austin, 2013).
20. A facility this size is capable of producing approximately 57 significant quantities of 93% enriched HEU annually, if it dedicated its entire capacity to the task.
21. Boyer, "Safeguards Approaches for Gas Centrifuge Enrichment Plants."
22. It is standard to keep 75 days worth of feed on-site; 84 is used here to simplify calculations (exactly three inspection cycles). D.M. Gordon et al., "An Approach to IAEA Material-Balance Verification at the Portsmouth Gas Centrifuge Enrichment Plant," *Proceedings of the Fifth Annual Symposium on Safeguards and Nuclear Material Management*, 1983.
23. C. Listner et al., "Quantifying Detection Probabilities for Proliferation Activities in Undeclared Facilities," (paper presented at Symposium on International Safeguards: Linking Strategy, Implementation and People, Vienna, Austria, 22–24 Oct 2014).
24. C. Bathke, "The Attractiveness of Materials in Advanced Nuclear Fuel Cycles for Various Proliferation and Theft Scenarios," (presented at International Workshop for Users of Proliferation Assessment Tools, Texas A&M University, 2009).
25. Note that this frequency of cascade hall inspections is likely not practical even with a resident inspectorate; the frequency was allowed to vary within wide bounds here to demonstrate the methodology.
26. J. N Cooley, "Progress in Evolving the State-Level Concept," (presented at the Seventh INMM/ ESARDA Joint Workshop: Future Directions for Nuclear Safeguards and Verification, Aix-en-Provence, France, 2011); M. Whitaker, M. Laughter, and D. Lockwood, "Information-Driven Inspections," (presented at the 2010 IAEA Symposium on International Safeguards, Vienna, Austria, 1–5 Nov. 2010); Kory W Budlong Sylvester, Joseph F Pilat, and Chantell L Murphy, "Developing State-Level Approaches Under the State-Level Concept," (paper presented at Symposium on International Safeguards: Linking Strategy, Implementation and People, Vienna, Austria, 1 Oct 2014).