Transactions of the Korean Nuclear Society Spring Meeting
Jeju, Korea, May 29-30, 2014

A Game Theoretic Approach to Nuclear Security Analysis against Insider Threat

Kyo-Nam Kim [a], So Young Kim [b], Erich Schneider [c], Man-Sung Yim [a]
[a]*Dept. of Nuclear and Quantum Engineering, Korea Advanced Institute of Science and Technology*
[b]*Graduate School of Science and Technology Policy, Korea Advanced Institute of Science and Technology*
[c]*Nuclear and Radiological Engr. Program, Dept. of Mechanical Engineering, the University of Texas at Austin, USA*
[*] *Corresponding author: charismak@kaist.ac.kr*

1. Introduction

Measures to advance nuclear safety and nuclear security are conventionally considered to serve distinct, and often conflicting, objectives. However, the Safety-Security-Safeguards (3S) interface is recently emerging as a key issue within both nuclear safety and security research. Insider threat to nuclear facilities is a particularly important issue in this regard, as it can compromise both safety and security of a nuclear installation. As individuals with authorized access to a facility and system who use their trusted position for unauthorized purposes, insiders are able to take advantage of their access rights and knowledge of a facility to bypass dedicated security measures [1]. They can also capitalize on their knowledge to exploit any vulnerabilities in safety-related systems, with cyber security of safety-critical information technology systems offering an important example of the 3S interface. Because insiders are capable of carrying out defeat methods not available to outsiders and have more opportunities to select the most vulnerable target and the best time to execute the malicious act, insider attacks are the key threat to the 3S interface.

This study examines a novel quantitative framework for performing nuclear security analysis against insider threat at a generic nuclear power plant. Most tools assessing the security threats focus on a limited number of attack pathways defined by the modeler and are based on probabilistic calculations. While this Probabilistic Risk Assessment (PRA) approach is appropriate for describing fundamentally random events like component failure of a safety system, it does not capture the adversary's intentions, nor does it account for adversarial response and adaptation to defensive investments [2,3]. To address these issues of intentionality and interactions, this study adopts a game theoretic approach. The interaction between defender and adversary is modeled as a two-person Stackelberg game. The optimal strategy of both players is found from the equilibrium of this game. A defender strategy consists of a set of design modifications and/or post-construction security upgrades. An attacker strategy involves selection of a target as well as a pathway to that target. In this study, application of the game theoretic approach is demonstrated using a simplified test case problem. The test case problem is based on a previous study done at University of Texas [2,3].

2. Model Description

For a test case problem, we model a simple nuclear facility with two targets - cooling tower and switchyard as shown in Fig. 1. Both are located inside the limited area of the facility [3].

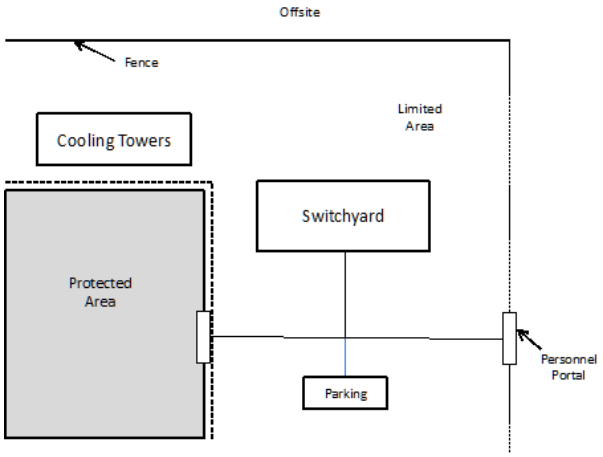

Fig 1. Conceptual depiction of the facility

With this conceptual design of the facility, networks or directed graph of arcs and nodes are modeled with nodes representing locations and arcs representing paths of movement between two locations. Each arc is assigned a non-detection probability and travel time.

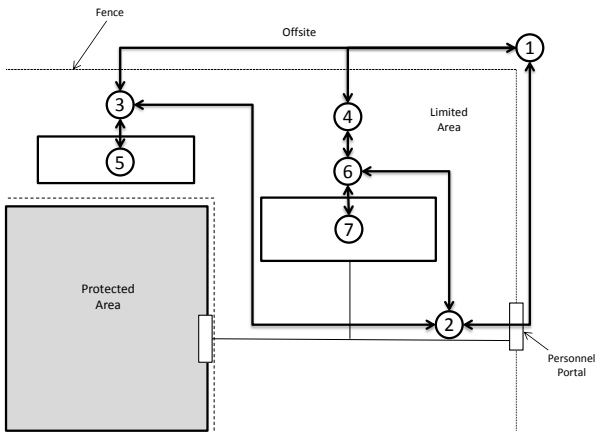

Fig 2. Network overlay of the facility

Note that the adversary succeeds not only by attacking the target but also by escaping the facility. We thus use a mirrored network in which the first half of the network contains pathways to the target and the second half (the reflection) contains paths of egress is suggested. Travel times and non-detection probabilities can differ on these two halves of the network, and the

ingress and egress networks themselves need not be perfectly symmetric.

The arc non-detection probabilities and travel times are populated by assessing the set of obstacles detecting or delaying an adversary traversing the arc and then estimating the probability of detection and time delay associated with each obstacle. It should be noted that the data in this model was constructed for use in student exercises for vulnerability analyst training, and is thus impractical to use in actual security analyses.

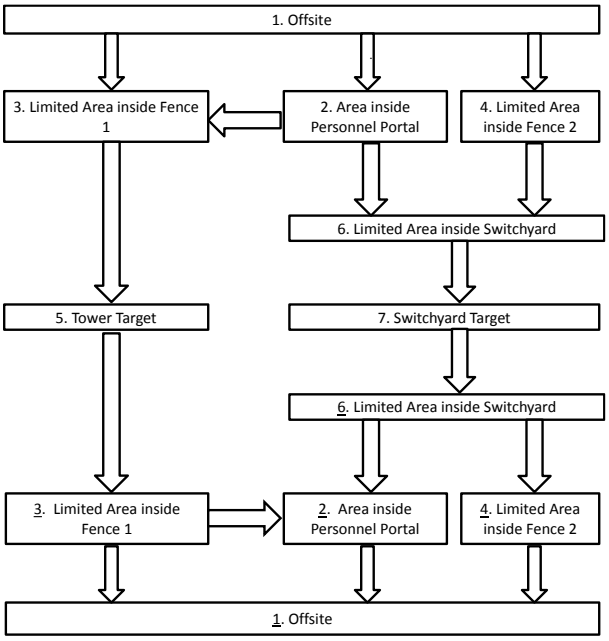

Fig 3. Mirrored network representation of facility

Table I. Obstacles location in baseline facility

Path	Path Type	Fence	Personnel Portal	Stationed guards	Random Searches	Non Guard Personnel	Roaming Guards	Alarmed Detection Device	Video Surveillance
1 → 2	Perso-nnel		1	1	1				1
2 → 1	Perso-nnel								1
1 ↔ 3	Fence	1				1	1	1	1
1 ↔ 4	Fence	1				1	1	1	1
2 ↔ 3	Prope-rty								1
2 ↔ 6	Prope-rty								1
3 ↔ 5	Attack Tower	1				1	1	1	1
4 ↔ 6	Prope-rty								1
6 ↔ 7	Attack Swit-chyard					2			1

Table II. Non-detection probabilities and travel times data

Path	Non-Detection Probability	Travel Time
1 → 2	0.86	35
2 → 1	0.97	10
1 ↔ 3	0.73	50
1 ↔ 4	0.73	50
2 ↔ 3	0.91	60
2 ↔ 6	0.91	60
3 ↔ 5	0.72	120
4 ↔ 6	0.91	60
6 ↔ 7	0.85	50

2.1. Model assumptions

While the previous research work [3] contains 6 model assumptions in greater detail, it does not consider the insider threat. We thus revise the model with additional assumptions as follows.

Insiders are defined as three group A, B, and C with three path concepts; Intrusion, Guidance, and Attack. Intrusion affects the outer path such as (1 → 2), (1 ↔ 3), and (1 ↔ 4). Guidance covers some paths that are inside the limited area like (2 ↔ 3), (2 ↔ 6), and (4 ↔ 6). Attack covers two paths which interact the target directly such as (3 ↔ 5) and (6 ↔ 7). Group A includes only Intrusion paths. Group B includes Intrusion and Guidance paths. Group C includes Guidance and Attack paths.

It is assumed that insider group can only affect the non-detection probability in this model. The non-detection probability was assumed to increase by 10% in Intrusion and Attack paths and 5% in Guidance paths.

2.2. Analytic framework

The two-person Stackelberg game is formulated by use of a mixed integer program (MIP). The actual model is constructed with the GAMS software program [4] that also returns expected consequence results.

2.3. Baseline problem

Based on the multi-target mirrored network and the MIP formulation, we define the following baseline problem with zero budget (B=0) for security upgrades. The baseline network was assumed to have the default security measures defined previously. Hence, even in

the absence of the security upgrades described below, an adversary is confronted with significant security measures. Numerical parameters in this section are assigned again for illustrative purposes.

2.4. Defender upgrades

The security upgrades include both design changes and security measures that may be installed after construction. This example assumes a single budget for all upgrades, yet it can be modified to separate the upgrades into security by design and operational security categories with their own budgets. Summary of security upgrades are shown below.

Table III. Summary of upgrade cost and effect

Upgrade ID	Cost	Impact
A	$$	Move tower inside of PIDAS (perimeter intrusion detection and assessment system): add a node and arc on either side of tower target (in series), each with non-detection probability of 70%.
B	$$$$	Build additional redundant cooling tower and switchyard, each with half the consequence. Adversary has option to attack both sides by traversing additional arcs in the system.
C	$$	Multiply travel time by 2 for arcs on either side of Tower target.
D	$$	Reduce non-detection probability by 20% for all arcs between nodes (2,3), (2,6), and (4,6).
E	$$$	Increase travel time by 25% for all arcs between nodes (1,3) and (1,4).
F	$	Reduce response time by 20 seconds.
G	$	Increase defeat probability from 0.8 (inside critical detection region 2, outside critical detection region 1) and 0.9 (outside critical detection region 2) to 0.9 and 0.95 respectively

3. Results and Discussion

Game-theoretic modeling is obviously inapplicable at a zero budget as there are no defender decisions to be made. With non-zero budgets, this model can be applied to choose the optimal security upgrades portfolio. Fig. 4 shows a summary of expected consequences as a function of budget increase from zero to fifteen for all groups. And Tables IV and V summarize the upgrades that are purchased at each budget level. At all groups, the expected consequence values are decreasing with the budget level. And the expected consequence values increase from those of No insider group to Group 1, 2, 3. Insiders marked as Group 3 have the largest capability to affect the non-detection probability. So the expected consequence of adversary is the highest at Group 3. Not only expected consequence but the solution of security upgrades also change in some budget level such as B=3, 4, 6, 7, 9, and 15. Security upgrades portfolio could also in consideration of the insider's assistance to adversary.

Table IV. Summary of results for each budget level; No insider group and group 1

	No insider		Group 1	
budget	Expected consequence	solution	Expected consequence	solution
0	0.962	None	1.038	None
1	0.887	G	0.965	G
2	0.619	A	0.671	A
3	0.592	A,G	0.637	A,F
4	0.546	A,F,G	0.595	B
5	0.471	B,G	0.508	B,G
6	0.441	A,B	0.475	A,B
7	0.406	A,B,G	0.442	A,B,G
8	0.318	A,B,D	0.340	A,B,D
9	0.278	A,B,D,G	0.301	A,B,D,G
10-12	0.235	A,B,D,F,G	0.254	A,B,D,F,G
13-15	0.220	A,B,D,E,F,G	0.237	A,B,D,E,F,G

Table V. Summary of results for each budget level; group 2and group 3

	Group 2		Group 3	
budget	Expected consequence	solution	Expected consequence	solution
0	1.080	None	1.155	None
1	1.009	G	1.093	G
2	0.699	A	0.752	A
3	0.676	A,F	0.726	A,F
4	0.607	A,D	0.640	A,D
5	0.550	B,G	0.598	A,D,G
6	0.508	A,D,F,G	0.544	A,D,F,G
7	0.454	B,D,G	0.523	B,D,G
8	0.365	A,B,D	0.448	A,B,D
9	0.310	A,B,D,F	0.376	A,B,D,F
10-12	0.265	A,B,D,F,G	0.335	A,B,D,F,G
13-14	0.265	A,B,D,E,F,G	0.335	A,B,D,E,F,G
15	0.265	A,B,C,D,E,F,G	0.335	A,B,C,D,E,F,G

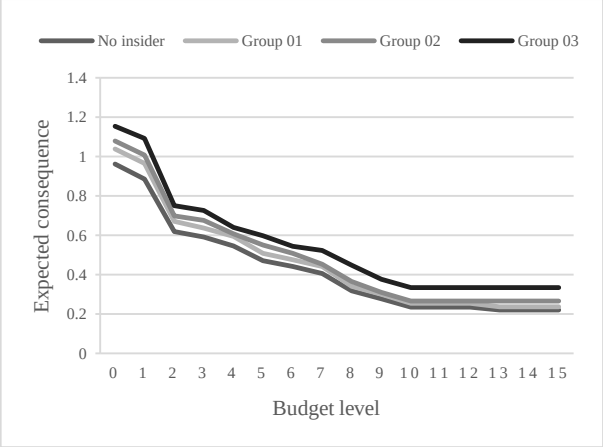

Fig. 4 Efficient frontier

4. Conclusion

This study constructs a game-theoretic model has been constructed for a physical protection system against attack by an intelligent adversary on a nuclear facility. Novel to our approach is the modeling of insider threat that affects the non-detection probability of an adversary. The game-theoretic approach has the advantage of modelling an intelligent adversary who has an intention and complete knowledge of the facility. In this study, we analyzed the expected adversarial path and security upgrades with a limited budget with insider threat modeled as increasing the non-detection probability. Our test case problem categorized three groups of adversary paths assisted by insiders and derived the largest insider threat in terms of the budget for security upgrades. Certainly more work needs to be done to incorporate complex dimensions of insider threats, which include but are not limited to: a more realistic mapping of insider threat, accounting for information asymmetry between the adversary, insiders, and defenders, and assignment of more pragmatic parameter values.

REFERENCES

[1] International Atomic Energy Agency, Preventive and Protective Measures against Insider Threats: Implementing Guide, IAEA Nuclear Security Series No. 8, IAEA, 2008.
[2] R. Ward, E. Schneider, A Game Theoretic Approach to Nuclear Safeguards Selection and Optimization, Dissertation at The Univ. of Texas at Austin, 2013
[3] B. Canion, C. Hadlock, A. Zolan, D. Morton, J. E. Bickel, E. Schneider, Game-theoretic Allocation of Security Investments at Nuclear Reactors, Risk Analysis, under review, 2013.
[4] R. E. Rosenthal, GAMS – A User's Guide, GAMS Development Co., 2014