Applying STAMP to Critical Infrastructure Protection

Joseph R. Laracy, *Student Member, IEEE* and Nancy G. Leveson, *Member, IEEE*

Abstract— **Classical risk-based or game theoretic security models rely on assumptions from reliability theory and rational expectations economics that are not applicable for security risks. Additionally, these models suffer from serious deficiencies when they are applied to software-intensive, complex engineering systems. Recent work in the area of system safety engineering has led to the development of a new accident model for system safety that acknowledges the dynamic complexity of accidents. System-Theoretic Accident Models and Processes (STAMP) applies principles from control theory to enforce constraints on hazards and thereby prevent accidents. Appreciating the similarities between safety and security while still acknowledging the differences, this paper introduces the use of STAMP to security problems. In particular, it is applied to identify and mitigate the threats that could emerge in critical infrastructure systems such as the air transportation network.**

***Index Terms*—Air transportation, Critical infrastructure, Public Safety, Security**

I. Introduction

"EVERY one of the whistleblowers interviewed by GAP (Government Accountability Project) warned that the airports are not safer now than before 9-11. The main difference is that life is now more miserable for the passengers." [1] This quote from a former red team leader at the FAA in 2003 is a powerful reminder of the inadequacies of the current state of security for the air transportation system. The ordinary traveler may have a perception of security as a result of airport inconveniences, but the determined terrorist can distinguish between security and its illusion. The interdisciplinary nature of the security problem is one of the key factors that make the solution so elusive. Traditional, disciplinary approaches on their own are often insufficient to accomplish the security goals of a complex system. Only a *comprehensive* methodology has the potential to succeed [2].

Many security related terms have moved into colloquial language and unfortunately lost their rigorous definitions. Key security terms are defined below to remove all ambiguity about the authors' use of these terms.

Manuscript received March 9, 2007. This work was supported in part by the NSF under Grant CNS-0550008.

J. R. Laracy is a doctoral student in the Complex Systems Research Laboratory at MIT, Cambridge, MA 02139 USA (laracy@mit.edu).

N. G. Leveson is a Professor and Director of the Complex Systems Research Laboratory at MIT, Cambridge, MA 02139 USA (leveson@mit.edu).

- **Security**: A *system* property that implies protection of the informational, operational, and physical elements from malicious intent.
- **Vulnerability**: A weakness in a system that can be exploited to violate the system's intended behavior relative to security.
- **Threat**: An intentional action aimed at exploiting a vulnerability.

Large scale, complex systems require *physical*, *information (communication* and *computer*), and *operational* security [3]. Vulnerabilities often emerge in an engineering system when one or more of the aforementioned domains are omitted. It is important to note that attackers rarely choose to directly engage the most secure aspects of a system such as the cryptographic algorithms. In the words of internet security expert, David Clark, "Encryption is perfect, no one break codes, they just steal the key." [4]

II. Limitations of Classical Approaches

A variety of approaches exist both in industrial practice and the academic literature for conducting security analyses on large infrastructure systems. These methods include "best practice engineering," quantitative risk assessment, game theory, and red teaming. The four classical approaches each have their own strengths and weakness but unfortunately do not provide total coverage for the *system security* problem.

The most common security technique is simply to apply best practices. This approach is usually conducted in an ad hoc way and reduces or removes only the most obvious vulnerabilities [5]. If a systematic approach is taken to develop a comprehensive body of best practice literature, the best practice approach would be far more useful to engineers. Usually, security experts will employ one or more of the following methods to supplement best practice approaches.

A. Risk Based Security

Risk-based security seeks to quantify security risks by assigning severity and likelihood ratings to attack scenarios. The emphasis of this technique has been on risk-based decision making. The goal is to direct security investments as opposed to modeling particular kinds of threats. The approach is derived from reliability models of accident causation that are rooted in a chain-of-events perspective. Whether part of a preliminary hazard analysis or an accident reconstruction activity, the reliability engineer attempts to understand the potential or actual accident by identifying the events or faults

that could initiate the accident. Such fault and event trees are usually part of a method called probabilistic risk assessment (PRA). PRA was developed in the mid 1970s to improve nuclear power plant safety [6]. A probabilistic risk assessment is a four step process:

1. Identify undesirable events.
2. Identify accident scenarios (sequences of events).
3. Estimate the probability of each scenario either based on statistical testing data, or expert judgment if scenarios are rare.
4. Rank the accident scenarios according to likelihood.

The framework yields a probability for each undesirable event identified in the first step.

PRA turned out to be very successful for assessing risks in nuclear power shut-down systems. Such systems were historically very simple, electro-mechanical systems designed to minimize unnecessary complexity and use proven analog electrical technologies. PRA carries with it a number of important assumptions:

1. The events or faults in the trees are collectively exhaustive — all possible events are identified.
2. The events or faults in the trees are mutually exclusive — they cannot occur simultaneously.
3. The probability of each scenario is accurate enough to be useful to decision makers.

However, when complex systems like the air traffic management system are considered, serious questions arise regarding the appropriateness of PRA. Recently, researchers in the field of PRA acknowledged that PRA should not be the sole basis for decision making and that the quantitative results should be part of risk-*informed*, not risk-*based* decisions. They acknowledge that human factors, software, safety culture, and design errors are not well handled by PRA [7].

Given the central role of human factors, software, culture, and design errors in security, PRA's applicability to security problems is also dubious [8]. Donn Parker makes an insightful observation in this regard:

> "Security risk is not measurable, because the frequencies and impacts of future incidents are mutually dependent variables with unknown mutual dependency under control of unknown and often irrational enemies with unknown skills, knowledge, resources, authority, motives, and objectives – operating from unknown locations at unknown future times with the possible intent of attacking known by untreated vulnerabilities that are known to the attackers but unknown to the defenders." [9]

Nonetheless, a variety of researchers have attempted to supplement pure, reliability-based PRA with other techniques to make it relevant to security. For example, Michaud and Apostolakis developed a scenario-based methodology to rank elements of an infrastructure system according to their value to the stakeholders. Through a combination of probabilistic risk assessment, multi-attribute utility theory, and graph theory, the methodology models the infrastructure system as a network. After scenarios are generated, a value tree is built to evaluate scenarios and their consequences. The value tree incorporates the disutility of each scenario and vulnerability categories are assigned a ranking ranging from level I (Red) to level V (Green). The high level goal of this approach is to answer the following questions:

> What can go wrong?
> What are its consequences?
> How likely is it? [10]

The first two questions are more effectively answered by qualitative hazard or threat analysis techniques, while the last question may not be answerable in rare events such as the terrorist attack of 9/11/2001.

The effect of misapplying quantitative, probabilistic techniques can lead to the dangerous illusion of strong security. Good work has been done by Dean Wilkening in missile defense strategy development comparing shoot-look-shoot and barrage firing options [11]. This research was based on extensive *empirical data* from test firing exercises and live military operations. Extensive research indicates that questions of likelihood for rare events cannot be accurately estimated and expert judgment is often systemically biased [12]. Simplifying assumptions such as assuming that terrorist groups will only plan one method of attack are inconsistent with reality. Furthermore, developing event or decision trees with mutually exclusive and collectively exhaustive attack scenarios can easily produce a tree that exceeds the intellectual manageability of the engineer. It is unlikely that reductionist, bottom-up approaches will succeed.

B. Game Theory

Bier [13] asserts that managing risks from intelligent adversaries is very different from other types of risk and suggests game theory over decision theory. Previous work in this area focused on "policy insights" such as the relative merits of deterrence and other protective measures [14]. Sandler and Arce present a number of compelling reasons for the applicability of game theory to security problems [15]:

However, game theory "requires strong assumptions about the availability of mutual information and the rationality of opponents." [16] As mentioned earlier, empirical work by Tversky and Kahneman [12] has shown that these assumptions often break down in reality. Additionally, traditional games are organized to pursue a minimax solution for a two-person, zero-sum game. However, as Banks and Anderson point out, such a model is only an approximation because defender and attacker will value successful and failed attacks differently [16].

Many game-theory models of security carry the traditional, simplifying assumption that the probability of a successful terrorist attack on a location is a convex function of the defensive resources. Some security measures, such as relocating a facility to a more secure location, are inherently discrete. Discretization introduces step changes into the function so there is no longer a smooth, convex function due to declining marginal returns on defensive investments. Also, if a particular level of defensive investment completely deters an attack, the probability of terrorist success drops rapidly

beyond that point. This scenario would also produce a non-convex function in certain regions. When non-convex functions are permitted, multiple local optima may emerge, thereby complicating the defense resource allocation problem [13].

One useful insight that a game-theoretic analysis can offer is problem framing. Palmore's work [17] on preventive defense against ballistic missile attack is an example of research that offers insights into problem formulation without the pitfalls of misapplying quantitative methods. Game theory's role in security focuses on analyses related to assessing strategies for how to allocate national antiterrorism expenditures, measuring how military strategies encourage or discourage terrorism, assessing insurance risks, and evaluating the effects of focusing either on deterrence or preemption [18]. As the list above indicates, game-theoretic models focus on strategic decision making. Questions of how to design and operate infrastructure systems that may be the target of terrorist attacks is the focus of the systems-theoretic analysis introduced in this paper.

C. Red Teaming

The words of Dr. William Schneider, Jr., Chairman of the Defense Science Board, best capture the state of red teaming: "Red teams can be a powerful tool to understand risks and increase options. However, the record of use of red teams in DoD is mixed at best." [19] Red teaming is an excellent activity to complement other security analyses and activities as well as reduce the complacency that often sets in after extended periods without attacks. The goal of any red team is to challenge the plans, programs, and assumptions of the client organization. Teams may challenge organizations at strategic, operational, or tactical levels depending on the area that needs the most attention.

The greatest benefit derived from red teaming exercises is "hedging against catastrophic surprises." A good red team is capable of elucidating a deeper understanding of an adversary's options, and identifying vulnerabilities in concepts, programs, plans, postures, and strategies. Red teams also challenge "the accepted assumptions and accepted solutions" as well as identify inexperience. They may function as surrogate adversaries, devil's advocates, or simply as sources of independent judgment.

Schneider also points out that "red teaming is important but it is not easy nor often done very well." He identifies the following causes of failure:

The red team:

1. Does not take its assignment seriously.
2. Could lose its independence.
3. Could be too removed from the decision making process.
4. Could have inadequate interaction with the "blue" (team) and be viewed as just another sideline critic.
5. Could destroy the integrity of the process and lose the confidence of decision makers by leaking its findings to outsiders.

Red team effectiveness is easily impaired by a corporate culture that does not value criticism and challenge. Managers that do not want issues to arise that may "rock the boat," dysfunctional interaction between red and blue teams, unqualified red team staff, and calling in a red team when the problem has grown out of control often does little to mitigate disaster. The red team must have independence with accountability as well as a process that enables the game results to be seriously considered by senior management [19].

Unfortunately, the red teaming process failed miserably before 9/11/2001. Testimony by Bogdan Dzakovic, an FAA red team veteran, to the National Commission on Terrorist Attacks Upon the United States, on May 22, 2003, reveals how a good red team can become completely ineffective in the face of management resistance. The Presidential Commission investigating the bombing of Pan Am 103 in 1990 created the FAA red teams that are in place today. After TWA 800's crash, Congress passed the FAA Reauthorization Act of 1996. The law states that "…the Administrator (of the FAA) shall conduct periodic and unannounced inspections of security systems of airports and air carriers to determine the effectiveness and vulnerabilities of such systems…" Later, in 1997, a White House Commission stated that "…Red Team testing should also be increased by the FAA, and incorporated as a regular part of airport security action plans. Frequent, sophisticated attempts by these Red Teams to find ways to dodge security measures are an important part of finding weaknesses in the system and anticipating what sophisticated adversaries of our nation might attempt."[1] Unfortunately, as Dzakovic's testimony indicates, the value of these red teams had been seriously undercut:

> "Although we breached security with ridiculous ease up to 90% of the time, the FAA suppressed these warnings. Instead we were ordered not to retest airports where we found particularly egregious vulnerabilities to see if the problems had been fixed. Finally, the agency started providing advance notification of when we would be conducting our "undercover" tests and what we would be checking." [1]

Given the limitations of traditional approaches, the goal of this paper is to introduce a systems-theoretic security model that does not rely on the assumptions of quantitative risk assessment, considers issues at a level closer to system design and operation compared to game theory, and supports successful red teaming.

III. STAMP-Sec

STAMP-Sec views security incidents as the result of inadequate control, rather than strictly a failure event, such as cracking a code or a fault in a cryptographic device [20]. Security is an emergent system property that is achieved through the enforcement of *constraints*. This perspective allows security problems to be transformed into control problems for which powerful intellectual tools can be employed. *Control structures* are defined to capture the communication and control in the system and illustrate the

presence and absence of essential feedback. They are hierarchal in nature and should be constructed both for system development and system operation.

Security must be designed into a system and be a conscious part of how it is operated. Historically, systems where security was added in "after the fact" have been plagued by systemic security risks. For example, current approaches to information security suffer from serious deficiencies as evidenced by the influence of SPAM, internet worms, viruses, phishing, and other attacks that plague the common internet user. This is largely a result of the fact that network research in the 1960s through the 1980s focused on achieving performance objectives with little emphasis on security. As a result, when threats began to emerge in the 1990s, internet security was approached from an ad-hoc perspective – applying patches to vulnerabilities already identified by attackers. The problem remains that the underlying architecture was not designed to support strong security.

A STAMP control structure informs design by defining the necessary communication and control between subsystems and components to enforce security constraints. There are many ways inadequate control can lead to a security system being compromised. STAMP provides a useful categorization scheme that captures most security control flaws. Broadly, they fall into one of three categories: Inadequate enforcement of constraints, inadequate execution of control actions, or inappropriate or missing feedback [21]. The introduction of a malicious agent does not violate the assumption of the taxonomy originally developed for safety. In a safety scenario, poor engineering or management may offer inadequate enforcement of constraints, execution of control actions, or feedback such that a hazard is "exploited" inadvertently in system operations. In a security scenario, poor engineering or management may offer inadequate enforcement of constraints, execution of control actions, or feedback such that a vulnerability is created that may be intentionally exploited in system operation. Whether one is concerned with safety or security, the problem is inadequate control. STAMP-Sec extends the safety list to capture security issues:

1. Inadequate Enforcement of Constraints (Control Actions)
    1.1. Unidentified threats
    1.2. Inappropriate, ineffective, or missing control actions for identified threats
        1.2.1. Design of control process does not enforce constraints
            1.2.1.1. Flaws in creation process
            1.2.1.2. Process changes without appropriate change in control (asynchronous evolution)
            1.2.1.3. Incorrect modification or adaptation
        1.2.2. Process models inconsistent, incomplete, or incorrect
            1.2.2.1. Flaws in creation process
            1.2.2.2. Flaws in updating process (asynchronous evolution)
            1.2.2.3. Time lags and measurement inaccuracies not accounted for
        1.2.3. Inadequate coordination among controllers and decision makers (boundary and overlap areas)
2. Inadequate Execution of Control Action
    2.1. Communication flaw
    2.2. Inadequate actuator operation
    2.3. Time lag
3. Inadequate or missing feedback
    3.1. Not provided in system/organizational design
    3.2. Communication flaw
    3.3. Time lag
    3.4. Inadequate detection mechanisms

The reader should take note that many of these inadequacies are not associated with simply an event-based risk. Rather, *flaws in communication and control* as well as *time lags* and *flaws in the design process* contribute to threats

Effective communication between levels of a control structure hierarchy is essential to successful system security. This concept is easily seen in an example describing a small subset of the problems present in the air transportation system on September of 2001. On 9/11, the FAA regulations, standards, and certifications were both inherently inadequate and poorly enforced, risk 1.2.1 and 2.2. Legal penalties were not administered to security companies or airports that were shown to be dysfunctional by red team attacks, risk 2.2. Operations reports and red team results were ignored by mid-level managers and the results were not shared with senior leadership, risk 2.1 and 2.2. The Department of Transportation did not perform much better, as a major feedback flaw was present, risk 3.2. Incident and change reports, security assessments, and whistle blowers were unable to successfully communicate concerns to DoT leadership. According to Red Team leader and whistleblower Bogdan Dzakovic:

> "I went to the Department of Transportation's OIG. This too proved to be a wasted effort. A senior official in the Inspector Generals Office actually explained to us that because of the political situation between the FAA and the IG's office, the IG couldn't take any action against the FAA." [1]

This is a clear instance of inadequate execution of a control action, risk 2.2. Additionally, Bogdan's visit to the GAO was also unsuccessful: "The GAO people we spoke to were extremely concerned about our revelations, but explained they have no authority to actually do anything." [1] Clearly, the organizational design did not provide for the necessary feedback, risk 3.1. In fact, there was essentially no security feedback from the actual operating process, risk 3.4. Inadequate communication and control is shown in **Figure 1** by the dashed lines.

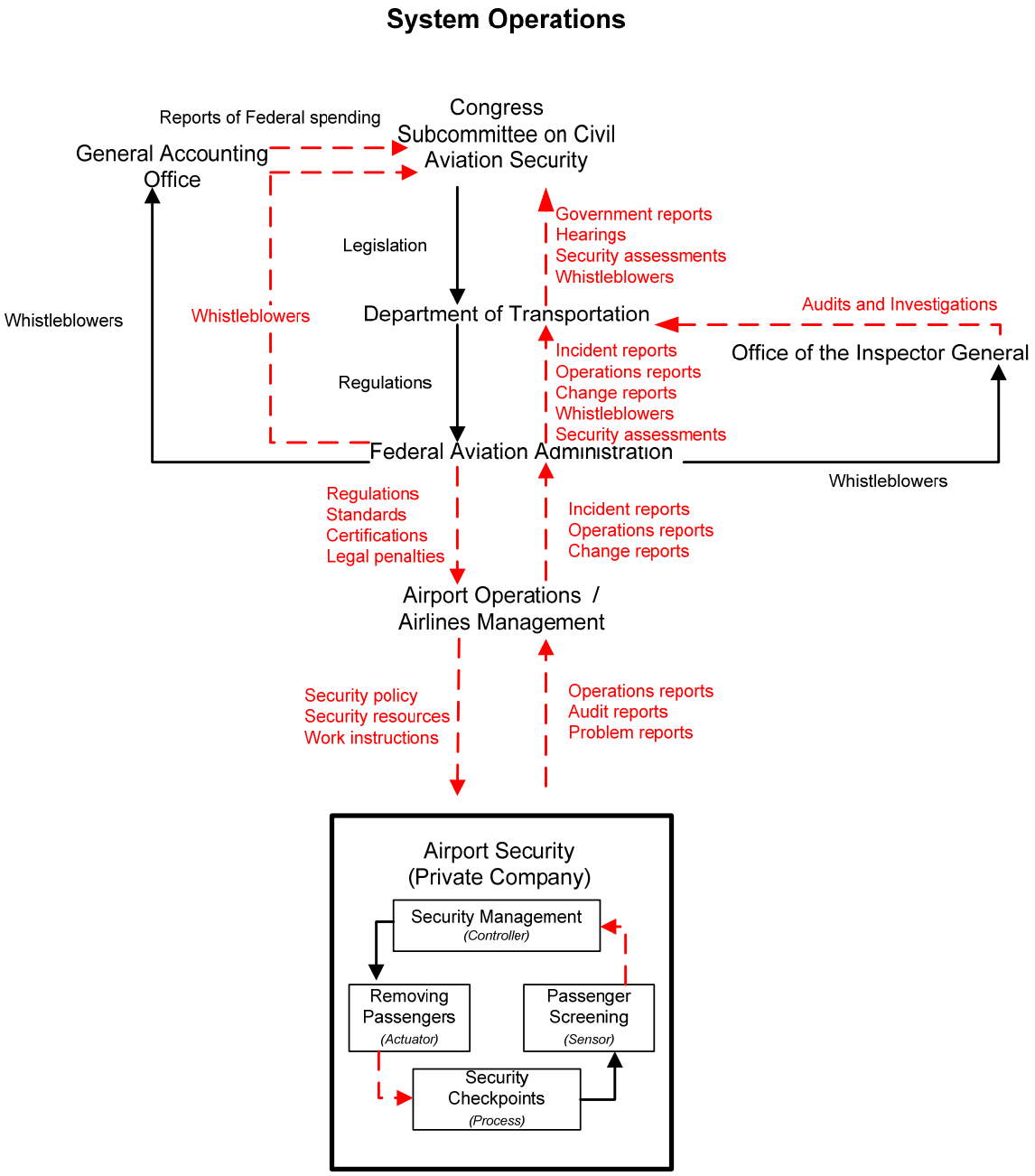

Figure 1 – A High Level Control Structure for Pre-9/11 Air Transportation System

In a top-down security engineering activity, threat analyses may be conducted using a variant of STPA, STamP-based Analysis [22]. The steps are provided below.

1. Identify the system-level threats.
2. Write security constraints for the threats.
3. Model the static control structure to prevent or mitigate the threats.
4. Assign constraints to the system components responsible for implementing them.
5. Define the control actions for the components that prevent or mitigate the threats and hazards.
6. Capture the behavioral dynamics with System Dynamics Modeling.

Threats are decomposed to the point where they can be rewritten as a design constraint. The complete list of constraints should be part of a system's requirements document. After that, the static control structure is modeled. Components in the control structure are assigned responsibility to execute the constraints. Finally, possible control actions for the components are defined [23]

A STAMP based analysis seeks to identify in the design phase both *how*, as show in the 9/11 example above, and *why* inadequate control could occur in a complex system. The *why*, or behavioral dynamics, is explained with a System Dynamics (SD) modeling [24]. SD is a modeling approach based on control theory and non-linear dynamics. The models themselves are systems of non-linear ordinary differential

equations that are solved numerically. As such, they contain state and rate variables to capture dynamic phenomena. A central modeling idea in SD is feedback loops; reinforcing and balancing loops shift state variables in counterintuitive ways. Without convincing simulation models, the intellectual manageability limitations cause humans to revert to linear thinking. Future work will demonstrate the utility of SD modeling in security analysis.

The system-level threats for an air transportation system are:

1. A terrorist takes control of or disrupts an aircraft or persons onboard.
2. A terrorist takes control of or impersonates air traffic control.
3. A terrorist sabotages an aircraft.
4. A terrorist shoots an aircraft down.
5. A terrorist disrupts the critical infrastructure of the air transportation system (i.e. destroy a runway or radar).
6. A terrorist interferes with the aircraft communication, navigation, or surveillance systems.

As an example, Threat 1 may be refined down to:

A terrorist attempts to make a "kamikaze" run on another aircraft or ground target. His TCAS is disabled.

This threat motivates a design constraint that does not permit TCAS to be disabled from within the cockpit. The constraint is enforced by the aircraft development contractor, governmental regulations, and the aircraft maintenance organization. The control action for the contractor would be to design the aircraft such that a crewmember or passenger could not disable TCAS mid-flight. The regulator control action is to create FAA regulations that mandate the design requirement. Maintenance organizations are also required not to make any modifications to the aircraft that would violate the regulation.

STAMP-Sec is a promising approach to security engineering for infrastructure systems because of its holistic, top-down approach. Systems engineers are adept at writing requirements so the security constraint is a logical extension to a design activity. No simplifying assumptions are part of the analysis and defining an exhaustive list of events or vulnerabilities does not need to be attempted. Also, the rationality of terrorists is not assumed. The effectiveness of STAMP control actions can be assessed with red teaming exercises. Red teams would also benefit from knowing the security constraints that the system under testing is supposed to enforce. Finally, expert knowledge currently being applied in ad-hoc approach can easily be integrated in a STAMP analysis.

I. Acknowledgments

The authors would like to thank Brandon D. Owens, Margaret Herring, and Dr. Nicolas Dulac for their thoughtful comments and suggestions.

References

[1] B. Dzakovic, "Statement of Bogdan Dzakovic to the National Commission on Terrorist Attacks Upon the United States," 2003.

[2] R. Anderson, *Security Engineering*. New York: Wiley Computer Publishing, 2001.

[3] J. R. Laracy, "A System-Theoretic Security Model for Large Scale, Complex Systems Applied to the Next Generation Air Transportatin System (NGATS)," in *Engineering Systems Division*. vol. Master of Science Cambridge: MIT, 2007.

[4] D. Clark, "Personal Communication on Security," J. Laracy, Ed. Cambridge, 2006.

[5] J. Laracy, "A Systems Theoretic Accident Model Applied to Biodefense," *Defense and Security Analysis*, vol. 22, pp. 301-310, September 2006.

[6] G. Apostolakis, "The Nuclear News Interview - Apostolakis: On PRA," in *Nuclear News*, 2000, pp. 27-31.

[7] G. E. Apostolakis, "How Useful is Quantitative Risk Assessment?," *Risk Analysis*, vol. 24, pp. 515-520, November 3, 2004.

[8] J. Laracy, "Addressing System Boundary Issues in Complex Socio-Technical Systems," in *Proceedings of the 5th Annual Conference on Systems Engineering Research* Hoboken, NJ, 2007.

[9] D. B. Parker, "Risks of Risk-Based Security," *Communications of the ACM*, vol. 50, March, 2007.

[10] D. Michaud, George E. Apostolakis, "Methodology for Ranking the Elements of Water-Supply Networks," *Journal of Infrastructure Systems*, vol. 12, pp. 230-242, 2006.

[11] D. Wilkening, "A Simple Model for Calculating Ballistic Missile Defense Effectiveness," in *Center for International Security and Cooperation*, 1998.

[12] A. Tversky, Daniel Kahneman, "Judgment under Uncertainty: Heuristics and Biases," *Science*, vol. 185, pp. 1124-1131, 1974.

[13] V. M. Bier, "Game-Theoretic and Reliability Methods in Counter-Terrorism and Security," in *Modern Statistical and Mathematical Methods in Reliability: Series on Quality, Reliability and Engineering Statistics*: World Scientific Publishing Co, 2005.

[14] B. S. Frey, S. Luechinger, "How to Fight Terrorism: Alternatives to Deterrence," *Defense and Peace Economics*, vol. 14, pp. 237-249, 2003.

[15] T. Sandler, Daniel G. Arce M., "Terrorism and Game Theory," *Simulation and Gaming*, vol. 34, pp. 317-337, September 2003.

[16] D. L. Banks, Steven Anderson, "Combining Game Theory and Risk Analysis in Counterterrorism: A Smallpox Example," in *Statistical Methods in Counterterrorism: Game Theory, Modeling, Syndromic Surveillance, and Biometric Authentication* A. G. Wilson, Gregory D. Wilson, David H. Olwell, Ed. New York: Springer, 2007.

[17] J. Palmore, Francoise Melese, "A Game Theory View of Preventive Defense Against Ballistic Missile Attack," *Defense and Security Analysis*, vol. 17, pp. 211-215, 2001

[18] R. D. Fricker, "Game Theory in an Age of Terrorism: How can Statisticians Contribute?," in *Statistical Methods in Counterterrorism: Game Theory, Modeling, Syndromic Surveillance, and Biometric Authentication* A. G. Wilson, Gregory D. Wilson, David H. Olwell, Ed.: Springer, 2005.

[19] W. Schneider, "The Role and Status of DoD Red Teaming Activities," in *Defense Science Board* September: Office of the Under Secretary of Defense for Acquisition, Technology, and Logistics, 2003.

[20] A. Rae, Colin Fidge, and Luke Wildman, "Fault Evaluation for Security-Critical Communications Devices," *Computer*, pp. 61-68, 2006.

[21] N. Leveson, "A New Accident Model for Engineering Safer Systems," *Safety Science*, vol. 42, p. 21, April 2004.

[22] N. Leveson, *System Safety Engineering: Back to the Future*. Cambridge, 2002.

[23] N. Leveson, "A New Approach to Hazard Analysis for Complex Systems," in *International Conference of the System Safety Society* Ottawa, 2003.

[24] N. Dulac, "A Framework for Dynamic Safety and Risk Management Modeling in Complex Engineering Systems," in *Aeronautics and Astronautics*. Ph.D. Thesis. Cambridge: MIT, 2007.