Journal of Strategic Security

Volume 4
Number 2 *Volume 4, No. 2, Summer 2011: Strategic Security in the Cyber Age*

Article 3

# Modeling Human Behavior to Anticipate Insider Attacks

Frank L. Greitzer , Ph.D.
*Pacific Northwest National Laboratory*, frank.greitzer@pnnl.gov

Ryan E. Hohimer
*Pacific Northwest National Laboratory*, ryan.hohimer@pnl.gov



pp. 25-48

**Recommended Citation**
Greitzer, Frank L. , Ph.D. and Hohimer, Ryan E.. "Modeling Human Behavior to Anticipate Insider Attacks." *Journal of Strategic Security* 4, no. 2 (2011) : 25-48.
DOI: http://dx.doi.org/10.5038/1944-0472.4.2.2
Available at: https://digitalcommons.usf.edu/jss/vol4/iss2/3

# Modeling Human Behavior to Anticipate Insider Attacks


## Abstract

The insider threat ranks among the most pressing cyber-security challenges that threaten government and industry information infrastructures. To date, no systematic methods have been developed that provide a complete and effective approach to prevent data leakage, espionage, and sabotage. Current practice is forensic in nature, relegating to the analyst the bulk of the responsibility to monitor, analyze, and correlate an overwhelming amount of data. We describe a predictive modeling framework that integrates a diverse set of data sources from the cyber domain, as well as inferred psychological/motivational factors that may underlie malicious insider exploits. This comprehensive threat assessment approach provides automated support for the detection of high-risk behavioral "triggers" to help focus the analyst's attention and inform the analysis. Designed to be domain-independent, the system may be applied to many different threat and warning analysis/sense-making problems.

Journal of Strategic Security
Volume IV Issue 2 2011, pp. 25-48
DOI: 10.5038/1944-0472.4.2.2

JSS
Published Quarterly by Henley-Putnam University Press
with support from the University of South Florida Libraries

# Modeling Human Behavior to Anticipate Insider Attacks

**Frank L. Greitzer**
**Ryan E. Hohimer**
*Pacific Northwest National Laboratory*
*Richland, WA USA*
*frank.greitzer@pnl.gov*

## Abstract

The insider threat ranks among the most pressing cyber-security challenges that threaten government and industry information infrastructures. To date, no systematic methods have been developed that provide a complete and effective approach to prevent data leakage, espionage, and sabotage. Current practice is forensic in nature, relegating to the analyst the bulk of the responsibility to monitor, analyze, and correlate an overwhelming amount of data. We describe a predictive modeling framework that integrates a diverse set of data sources from the cyber domain, as well as inferred psychological/motivational factors that may underlie malicious insider exploits. This comprehensive threat assessment approach provides automated support for the detection of high-risk behavioral "triggers" to help focus the analyst's attention and inform the analysis. Designed to be domain-independent, the system may be applied to many different threat and warning analysis/sense-making problems.

## Introduction

Imagine this (very general) scenario:

*John has been a productive employee for several years, but is extremely disappointed when he feels that other coworkers have taken credit for some of his accomplishments and he is passed over for a coveted promo-*

*tion. Filled with bitterness and frustration after being accused of inappropriate conduct at work, he negotiates with an outside entity to exploit his position to the benefit of the competition, planning later to join the competitor's organization.*

This brief scenario is a high-level description of a typical insider threat case. The insider threat refers to harmful acts that trusted insiders might carry out, such as something that causes harm to the organization or an unauthorized act that benefits the individual. Information "leakage," espionage, and sabotage involving computers and computer networks are the most notable examples of insider threats, and these acts are among the most pressing cyber-security challenges that threaten government and private-sector information infrastructures. The insider threat is manifested when human behaviors depart from established policies, regardless of whether they result from malice, disregard, or ignorance.

In the scenario above, if we jump back to the time of the Revolutionary War, we can see close parallels to the case of Benedict Arnold, who in 1780 conspired with the British to work towards the surrender of West Point following events between 1777 and 1779 involving his being passed over for promotion and being accused of financial schemes. Viewing the general scenario in more modern times, one can see parallels with the career of Aldrich Ames, a CIA operative from the late 1950s to the late 1980s. Ames initially received enthusiastic and positive reviews, but had continuing problems with alcoholism, security violations leading to reprimands, extramarital affairs that violated policy, and financial problems that reportedly led him to become a spy for the Soviet Union. In even more contemporary times, we may consider the case of accused WikiLeaks insider Bradley Manning, a despondent and disillusioned Army intelligence officer who experienced a series of emotional upheavals, including the breakup of a personal relationship, and whose disgruntled and inappropriate workplace behavior led to his demotion to Private/First Class before he allegedly leaked hundreds of thousands of U.S. Department of Defense and Department of State diplomatic cables.[1]

Surveys and studies conducted over the last decade and a half have consistently shown the critical nature of the problem in both government and private sectors. A 1997 Department of Defense (DoD) Inspector General report found that 87% of identified intruders into DoD information systems were either employees or others internal to the organization.[2] The annual e-Crime Watch Survey conducted by *Chief Security Officer (CSO) Magazine* in conjunction with other institutions reveals that for both the government and commercial sectors,[3] the most costly or damaging cybercrime attacks were caused by insiders, such as current or former employ-

ees and contractors. A recent report covering over one hundred forty-three million data records collected by Verizon and the U.S. Secret Service analyzed a set of one hundred forty-one confirmed breach cases in 2009 and found that 46% of data breaches were attributed to the work of insiders.[4] Of these, 90% were the result of deliberate, malicious acts; six percent were attributed to inappropriate actions, such as policy violations and other questionable behavior, and four percent to unintentional acts.

One might legitimately ask: Can we pick up the trail *before* the fact, providing time to intervene and *prevent* an insider attack? *Why is this so hard?* There are several reasons why development and deployment of approaches to addressing insider threat, particularly proactive approaches, are so challenging: (a) the lack of sufficient real-world data that has "ground truth" enabling adequate scientific verification and validation of proposed solutions; (b) the difficulty in distinguishing between malicious insider behavior and what can be described as normal or legitimate behavior (since we generally don't have a good understanding of normal versus anomalous behaviors and how these manifest themselves in the data); (c) the potential quantity of data, and the resultant number of "associations" or relationships that may emerge produce enormous scalability challenges; and (d) despite ample evidence suggesting that in a preponderance of cases, the perpetrator exhibited observable "concerning behaviors" in advance of the exploit, there has been almost no attempt to address such human factors by researchers and developers of technologies/tools to support insider threat analysis.[5]

Both the similarities and differences in cases throughout history reveal challenges for efforts to combat and predict insider threats. While the human factor has remained somewhat constant, the methods and skills that apply to insider exploits have changed drastically in the last few decades. In the time of Benedict Arnold, and even up to the time of the exploits of Aldrich Ames and another notorious insider, Robert Hanssen, an insider had to possess requisite knowledge, direct access to the information to be leaked, physical access to the recipient of the information, and a physical copy of the information to be exfiltrated. Compared to the WikiLeaks case, even twenty years ago it would have been necessary to use an 18-wheeler truck to transport the several hundred thousand documents involved in the WikiLeaks case. Today, insider crime does not even require specific knowledge of the information to be leaked, and gigabytes or more of information can be exfiltrated using various means, including thumb drives, email, and other modern information technology tools. Attribution is hard, and the ability to predict or catch a perpetrator in the act is severely limited, especially if the only means of detection is driven by workstation and network monitoring. Indeed, we have suggested that

the only way to be proactive is for the insider threat warning/analysis system to take non-IT "behavioral" or psychosocial data into account in order to capitalize on signs and precursors of the malicious activity that are often evident in "concerning behaviors" prior to the execution of the crime.[6]

In this regard, research suggests that in a significant number of cases, the malicious intent of the perpetrator was "observable" prior to the insider exploit. For example, a study by the Computer Emergency Response Team (CERT) Insider Threat Center,[7] a federally-funded research and development entity at Carnegie Mellon University's Software Engineering Institute, shows that 27% of insiders had come to the attention of either a supervisor or coworker for some concerning behavior prior to the incident. Examples of concerning behaviors include increasing complaints to supervisors regarding salary, increased cell phone use at the office, refusal to work with new supervisors, increased outbursts directed at coworkers, and isolation from coworkers.[8] As described in a recent article on rogue insiders, the extensive and ongoing investigation of insider threat by CERT has determined that most cases carry a distinct pattern. According to CERT's technical manager Dawn Cappelli, "Usually the employees either have announced their resignation or have been formally reprimanded, demoted, or fired."[9] In such cases, the article continues, the Human Resources department is aware of these high-risk personnel. The malefactors typically may be categorized as falling into one of two groups: either those who are moving to a new job and want to take their work with them, or those who are part of a well-coordinated spy ring operating for the benefit of a foreign government or organization.

## Goal of Insider Threat Research

In an operational context, security analysts must review and interpret a huge amount of data to draw conclusions about possible suspicious behaviors that indicate policy violations or other potentially malicious activities. They apply their domain knowledge to perceive and recognize patterns within the data. The domain knowledge that analysts possess facilitates the process of identifying the relevance of and connections among the data. In our examination of current practice by security, cyber security, and counterintelligence analysts, we have observed that typically the analyst uses a number of tools that monitor different types of data to provide alerts or reports about suspicious activities. This is primarily done in a forensic mode and within certain domains of data, such as output from Security Event and Incident Management (SEIM) systems, network and workstation/system log reports, web-monitoring tools,

access-control monitoring tools, and data loss/data leak protection (DLP) tools.[10] While these tools provide varying levels of protection, they are primarily forensic in nature, and in general the analyst has the critical and difficult responsibility for data fusion that integrates the analysis and "sense-making" across these disparate domains.

To date, no systematic methods have been developed that provide a complete and effective solution to the insider threat. Our goal is to create, adapt, and apply technology to the insider threat problem by incorporating into a reasoning system the capability to integrate different types of information that provide a useful picture of a person's motivation, as well as the capability and opportunity to carry out the crime. In this general context, our specific objective is to detect anomalous behaviors (insider threat indicators) before or shortly after the initiation of a malicious exploit.

Insider-threat assessment falls in the class of problems referred to as ill-structured, ill-defined, and wicked. Rittel and Webber defined "wicked problems" as those having goals that are incomplete, changing, and occasionally, conflicting.[11] Klein suggests that most real-world problems are not well-specified and do not involve "explicit knowledge."[12] Wicked problems defy clarifying goals at the start; we need to reassess our original understanding of the goals, and goals become clearer as we learn more. Methodologies are heuristic, involving discovery and learning through an iterative process. Thus, the objectives, concepts of operations, requirements, and other dictates in the proposed research and development are subject to periodic changes and maturation as the course of insider threat algorithms and software development matures.

The neocortex was the *inspirational metaphor* for the design of our reasoning framework, called CHAMPION (for **C**olumnar **H**ierarchical **A**uto-associative **M**emory **P**rocessing **I**n **O**ntological **N**etworks). The neocortex metaphor serves as inspiration for a functional (not structural) design that adopts functional requirements, as shown in Figure 1.

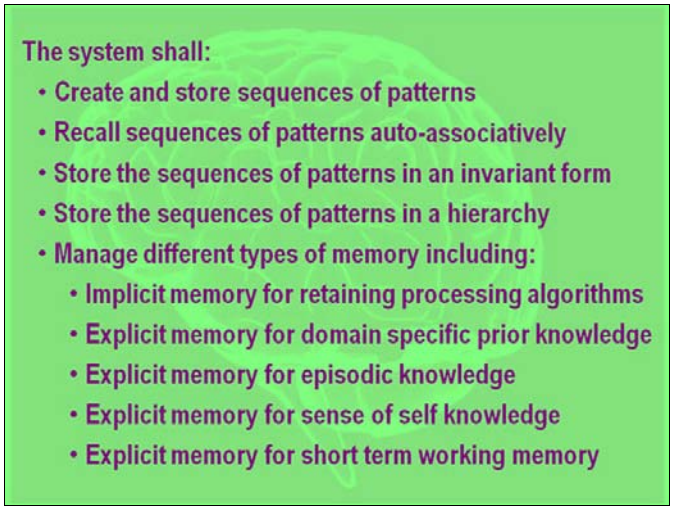


**Figure 1:** Functional requirements inspired by neocortex metaphor. The processing unit of the neocortex is the cortical column. For the CHAMPION reasoning system, the central processing unit is the Auto-associative Memory Component (AMC), which mimics the functionality of the cortical column.

## Multiple Domains of Data

The insider threat problem manifests itself within a *socio-technical system*, a combination of social, behavioral, and technical factors that interact in complex ways.[13] Knowledge within each of the factors is captured (or modeled) in a domain-specific ontology. This modeling approach organizes the notional concepts within a specific domain into a hierarchical mapping of those concepts. The ontologies are used by the computational reasoning system to interpret patterns within the data.

The rationale for our approach of integrating across different domains of data is based on a body of scientific research and case studies in the field of insider threat, cyber security, and social/behavioral sciences, from which it has been widely concluded that behavioral and psychosocial indicators of threat risk should be taken into account by insider threat analysis systems.[14] Indeed, Gudaitis and Schultz have separately argued that integrated solutions are required.[15, 16] As Schultz observed, there is a need for a "new framework" for insider threat detection based on multiple indicators that not only address workstation and network activity logs but

also include preparatory behavior and verbal behavior, among others. Thus, analysis of workstation and network data is a necessary, but not sufficient, condition for proactive insider threat analysis.

A recent review describes many technical approaches to intrusion detection (including insider threats) that may be characterized according to the categorization of techniques in terms of threshold, anomaly, rule-based, and model-based methods.[17, 18] Threshold detection is essentially summary statistics (such as counting events and setting off an alarm when a threshold is exceeded). Anomaly detection is based on identifying events or behaviors that are statistical outliers; a major drawback of this approach is its inability to effectively combat the strategy of insiders to work below the statistical threshold of tolerance and, over time, train systems to recognize increasingly abnormal behavior patterns as normal. Rule- or signature-based methods are limited to work within the bounds of the defined signature database; variations of known signatures are easily created to thwart such misuse-detectors, and completely novel attacks will nearly always be missed. Model-based methods seek to recognize attack scenarios at a higher level of abstraction than the other approaches, which largely focus on audit records exclusively as data sources. A sample of data sources for host- and network-based monitoring data that are relevant to insider threat detection is shown in Table 1.

**Table 1:** Representative Host/Network Cyber Data Monitored for Insider Threat Analysis

| | |
|---|---|
| • Registry entries | • File permissions |
| • Intrusion Detection System (IDS) events | • Access to account |
| • Firewall logs | • Email content capture |
| • Domain Name Server (DNS) logs/Internet sites accessed | • Email headers |
| • Host event logs | • Instant messaging |
| • Host print logs | • Keystroke records |
| • Network print logs | • Digital signatures |
| • Database server logs | • Local stored or cached files |
| • Web server logs | • Proximity card data |
| • Search engine queries (from query logs) | • Applications installed, patch status, version numbers for host computer |
| • Known software signature | |

Tools that focus on behaviors represent a significant advancement, but strict adherence to a statistically measured anomaly detection approach (which is in common use today) allows for gaps that may be exploited and/or go undetected. At the same time, tools that focus on policy violation assessment provide an effective first line of defense but allow for insider exploits that avoid policy violations. A more integrated approach that combines these functions would represent a modest advancement. Still better performance should be achieved by incorporating model-based technology into an integrated analysis that includes more explicit "human factors" dimensions, as we describe next.

In keeping with an approach that attempts to reflect relationships between certain personality characteristics and counterproductive work behavior or higher-risk employees, we conducted discussions with human resources (HR) professionals and managers at our organization to identify behavioral "proxies" for such characteristics that may, to varying degrees, produce a heightened concern about possible insider threat risks. Informed by the Five Factor Model (FFM) and previous research and case studies documenting personality disorders and factors of concern,[19] these discussions focused on the kinds of behaviors that would likely be observed and "known" by managers and HR staff because of the level of concern that they bring about. The model that evolved from these discussions was therefore highly observation-based, i.e., focusing on observable behaviors that could be recorded and audited. Therefore, although the model is based on behavioral observables, it can support making inferences about the possible psychological, personality, and social state of an employee; hence we refer to our model as a "psychosocial" model to capture the wide spectrum of inferences it is capable of producing.

The implementation of psychosocial reasoning used a data-driven approach based on personnel data that were likely to be available.[20] The indicators used in the model, such as disgruntlement, anger management issues, and disregard for authority, are shown in Figure 2. It is worth noting that these psychosocial indicators contribute differentially to the judged level of psychosocial risk, with disgruntlement, difficulty accepting feedback, anger management issues, disengagement, and disregard for authority having higher weights than other indicators, for example.



https://digitalcommons.usf.edu/jss/vol4/iss2/3
DOI: http://dx.doi.org/10.5038/1944-0472.4.2.2

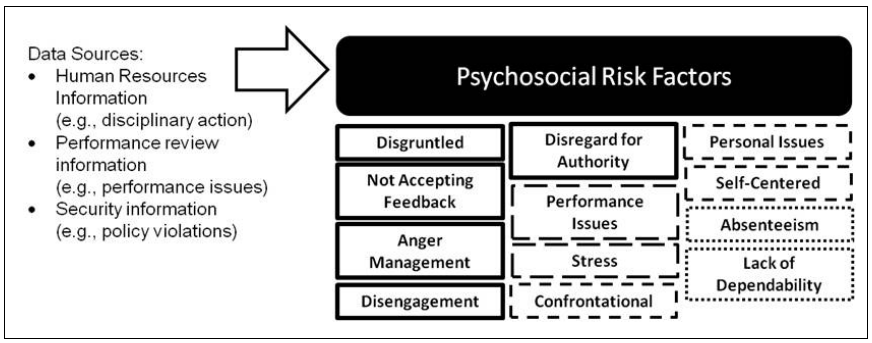


**Figure 2:** Psychosocial Indicators. Based on our preliminary research, factors with solid/heavy outlines are considered more important than those with lighter outlines in determining insider threat risk.

## Modeling Framework

A basic premise is the separation of the domain knowledge from the reasoning framework. If domain knowledge is hardcoded within the reasoning framework, then the framework's source code must be changed and recompiled whenever domain knowledge is updated. Equally important is the fact that this separation of domain knowledge from the reasoning framework maintains the domain-agnostic quality of the system, which enables its application to diverse problems without modification to the reasoning framework. We use the Ontology Web Language (OWL) as our knowledge representation language,[21] which implements the ontologies and knowledge bases of the system.

The main components of the CHAMPION system are:

- *Ontologies*, used for representing the specialized domain knowledge necessary to reason about the data. They provide the data-typing mechanisms.
- *Reifiers*, used for the ingesting of the primitive data types that are specified in the domain ontologies.
- *Memory*, used to store the facts asserted from the primitive data and the facts inferred by the reasoning system.
- *AMC (reasoning components)*, used to interpret the data assertions and infer new assertions.

The analysis of real-world data presents a challenge to computationally analyze very large graphs. The difficulty is not so much a *data reduction* problem as it is a *data interpretation* problem. If we think of a reasoning framework as a graph structure, a traditional approach builds the graph as part of the knowledge engineering process and then, when applied, the system conducts reasoning over the entire graph. In contrast, the CHAMPION hierarchy of reasoners comprises a "stack" of individual AMCs such that lower-level AMCs feed higher-level AMCs. The graph structure is built as data are analyzed; this produces a dynamic belief propagation network that takes in primitive data and pushes the interpretation of that data up the hierarchy. We can think of this as interpreting the current structure in the data and simplifying it with abstracting semantics. Just as we can stack the AMCs, we can stack collections (*regions*) of AMCs that address reasoning or pattern recognition for different domains. Just as the AMC is analogous to the cortical column, the AMC region is analogous to a cortical region. Similarly, even higher-level collections of AMCs enable reasoning across such regions, providing a natural mechanism for high-level information fusion and analysis that is typically lacking in conventional monitoring/detection systems.

The innovation of using a hierarchical framework of reasoners allows us to constrain the requirements of each reasoner to a narrowly-defined purpose. We apply a *semantic layer* upon the data to enable graph-theoretic approaches for prediction, detection, and mitigation options. With a well-formed semantic layer, we can overcome computational intractability by performing reasoning on *subsets* of the semantic graph of data: Rather than implementing a monolithic reasoner that is required to reason over all the concepts represented in the entire semantic graph, each reasoner in the hierarchy is only required to reason about a small set of relevant concepts.

The belief propagation network performs a transformation of the low-level literal inputs into higher-level abstractions. As shown in Figure 3, the CHAMPION system takes in inputs from various sources and reasons about them. Ingesting and properly formatting the input data for a given domain is performed by a *reifier*, which instantiates the input from a data source and packages the information into an OWL representation called an *individual*. In turn these individuals are instantiated in java objects called *abstractions*. The *abstractions* are added to the *Working Memory* of the CHAMPION system.

Reifiers are responsible for asserting OWL *individuals* (primitives) into the Working Memory via *abstractions*. Although great care has been taken to make sure that AMCs are domain-agnostic, it is not possible to

keep the reifiers domain-agnostic. The reifier takes in raw literal data and forms an OWL *individual* that is defined by the domain ontology. When raw data needs to be reified, specific code is required to convert the raw data into a data-type defined in the domain ontology.

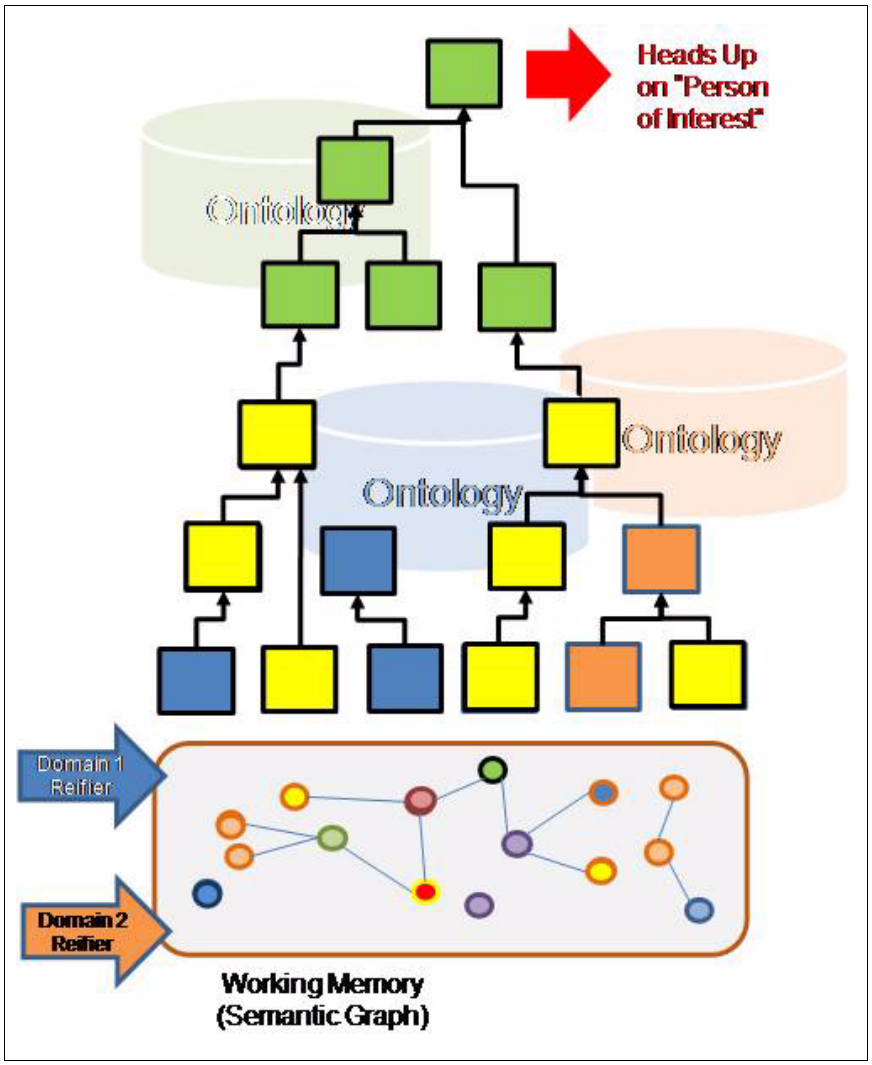


**Figure 3:** Functional Layout of CHAMPION Reasoning Process

For example, consider a reifier that is responsible for sensing information from a computer's security event log. The computer system logs security events. The responsibility of the reifier is to parse the security event logs and instantiate OWL *individuals*, wrap them in an *abstraction*, and add the abstraction into the Working Memory. The reifier must parse each line of the security event log and instantiate a single SecurityEvent *individual*, where the SecurityEvent is a class defined in the domain ontology. This *abstraction* is then inserted into the Working Memory. The AMCs (organized into various hierarchies, as indicated by the colored squares in Figure 3) further classify the SecurityEvent *abstraction*. Figure 4 illustrates this process in more detail. Here, an AMC focusing on failed login attempts checks the state of the subscribed SecurityEvent to determine if it matches a hypothesized pattern (in this case the main requirement would be that its event ID has a value of 529); and if a match occurs, this even may be typed as a FailedLogin, and a "FailedLogin" assertion is stored in Working Memory.

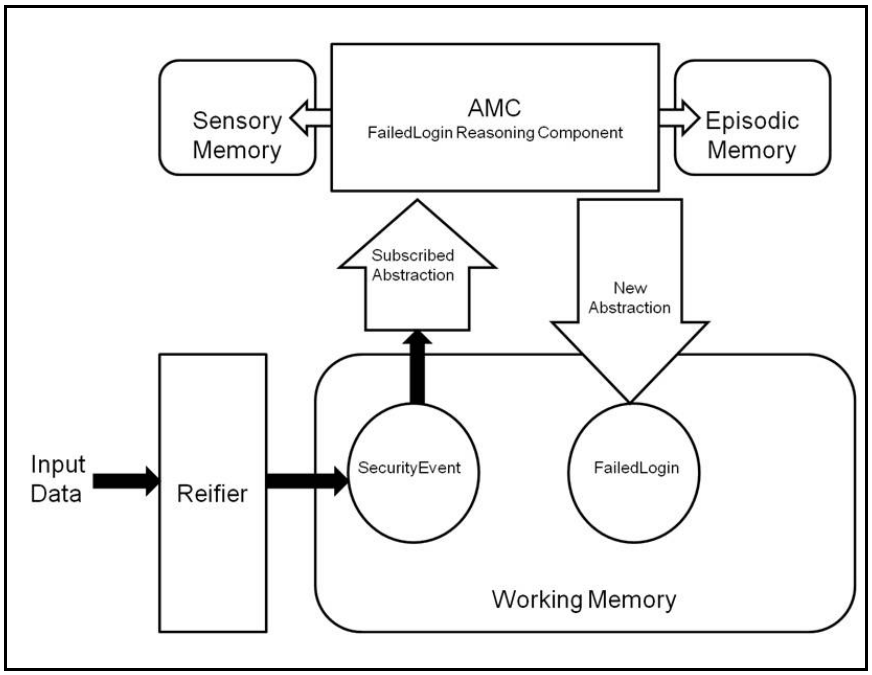


**Figure 4:** Illustration of process in which AMC subscribes to individuals deposited in Working Memory, analyzes the data, and deposits New Abstraction into Working Memory.

From a high-level perspective, the AMC subscribes to abstractions from Working Memory and publishes (infers) new abstractions back to Working Memory. In the CHAMPION system, the term "memory" refers to a machine-readable Knowledge Base (KB). A machine-readable KB stores knowledge in a form that can be processed by an inference engine to accomplish deductive reasoning about new knowledge inferable from current knowledge. A KB constructed with OWL is a collection of *individuals*, or instantiations of the class types defined in the domain ontology, which take the form of a semantic graph. A semantic graph comprises a set of nodes (the *individuals*) and the associated relationships (properties) among them. Therefore, we define a memory as a semantic graph of an *individual*, or *individuals*, in OWL format. There are several KBs in the CHAMPION system:

- Working Memory (also known as Short-Term Memory) stores the system's growing semantic graph
- Sensory Memory (localized memory in the AMC) stores the current inputs being reasoned over by the AMC between input cycles
- Episodic Memory (the memories of experience) stores the case library of the AMC and localized memory in the AMC

The process performed by the AMC within the belief propagation network of reasoners is shown in Figure 5. The system takes in data with no semantic tagging. The domain knowledge expressed in the domain ontology provides the data typing (semantic layer) which allows for the semantic tagging of the primitive data types being input into the system via the reification process. Once the "literal" facts are in the graph, the AMCs begin the process of further abstraction. New higher-level facts are added to the graph when the salient facts are present. The individual AMCs in the belief propagation network execute this process on the small subgraphs of the larger semantic graph.

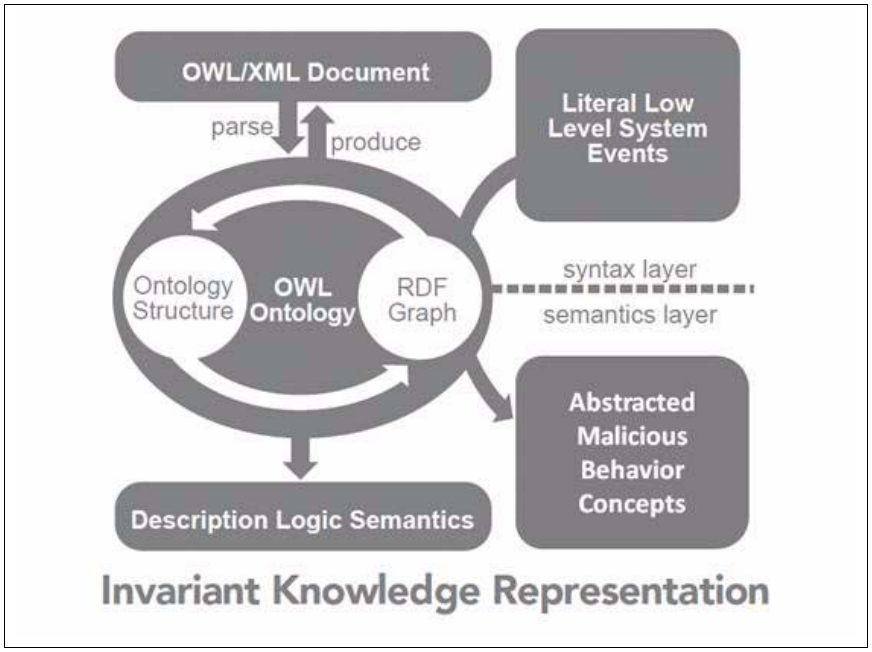


**Figure 5:** Recognition of Inputs and Adding Their Abstractions

Within each AMC there is a modified Knowledge Intensive Case-Based Reasoning (KI-CBR) structure.[22] Unlike traditional Case-Based Reasoning (CBR) systems that compare the current case under consideration with each case retained in their library of base cases,[23] this modified *knowledge-intensive* CBR approach uses formal ontological technologies to reason about the current case.[24] The reasoning approach uses Semantic Web Rule Language (SWRL) expressions.[25] If the SWRL rules can successfully construct a logically consistent abstraction from the current case, the case is added to the library of base cases and asserted back into the large semantic graph. This methodology offers two opportunities for the system to "learn:" First, the operator of the belief propagation network can provide feedback in the form of edits to the SWRL rule expressions that the reasoner uses to recognize abstractions. This is a direct editing of the description logics used to define the abstractions.[26] Second, once a sufficient number of cases have been added to the library of base cases, statistical analysis of the library can lead to SWRL rule modifications that can better classify future cases.

This implementation is quite different from the classical approach to CBR. The classical approach to CBR is to retrieve cases similar to the current problem from a case library. Then each case is reviewed to see if it is a viable solution to the problem. If it is, apply the solution and go on to the next problem. If there are no viable solutions in the library, the most similar case is revised to be a viable solution. If the new solution works, the case is retained in the library for future use.

## Operational Perspective

The "big picture" of the overall reasoning process implemented within the insider threat application of the CHAMPION system is shown in Figure 6. As described originally by Greitzer and colleagues, we envision the process of model-based reasoning that analyzes patterns of activity from data to observations to indicators to behaviors.[27]

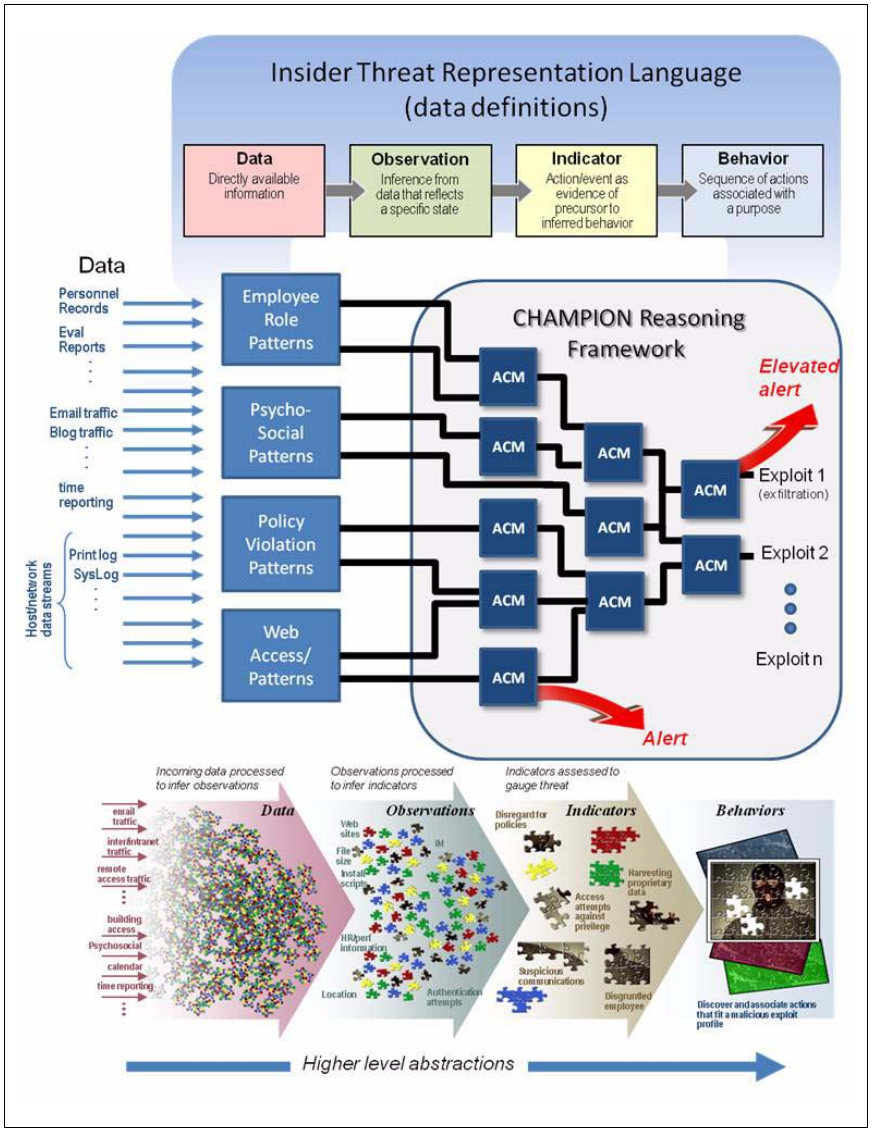


**Figure 6:** Conceptual model for reasoning framework conveys a sense-making task as assembling a complex puzzle through an abstraction process (data➔observation➔indicator➔behavior). Middle portion of the figure portrays the hierarchical system of reasoners that implements the abstraction process.



https://digitalcommons.usf.edu/jss/vol4/iss2/3
DOI: http://dx.doi.org/10.5038/1944-0472.4.2.2

In an operational form, the CHAMPION system ingests (reifies) data from a variety of data sources. These sources include data from SEIM systems, IDS systems, DLP systems, packet tracking systems, HR systems, etc. Each of the systems providing input into the CHAMPION reasoning system has an associated domain ontology that provides the data model and typing for that domain. Stated another way, data from specialized domains have their own AMC regions in the framework. "Bridging" domain ontologies defining the correlations between concepts in different domains are also present. For example, suppose there are three domain ontologies: (a) a cyber-activity ontology, (b) a social-networking ontology, and (c) a document-leakage ontology. An example of a concept defined in the cyber-activity ontology would be a staff member performs an *after-hours large download* (detection of this action comes from SEIM input). An example of a concept in the social-networking ontology would be a *large upload to social media*. In the document-leakage ontology, we could represent the concept of *suspected document leak* that could be defined as a combination of the *after-hours large download* and the *large upload to social media*. With even higher-level "abstractions" encoded in the knowledge representations, the system recognizes any of a number of exploits as equivalent (a print job is similar to a screen capture, storage of files on portable media, or uploading files to social media, etc.—all represent possible methods of exfiltrating data). The reasoning framework is therefore "set" to recognize activities that are consistent with patterns defined lower in the reasoning hierarchy, as well as higher-order patterns that bridge across multiple domains of data. This provides a more sophisticated approach to monitoring and threat analysis than is typical in current practice by helping the analyst correlate data over space and time and across varied data sources. This continual monitoring and analysis to recognize these types of integrated patterns serves to decrease the cognitive load on the analyst and focus her attention on possible threats that present the most critical security risks.

## Discussion

We have described the insider threat problem and a particular approach to proactively addressing the threat using an advanced belief propagation system that incorporates a variety of assessment data types. Two points are important to convey regarding this approach. First, it is useful to consider the problem addressed in more general terms, such as "problems in the workplace" or nonproductive behaviors in the workplace, and even issues of workplace violence. In particular, the insider threat problem is not limited to information technology attacks: consider the events of November 5, 2009, at Ft. Hood, Texas, where the accused perpetrator,

Army psychiatrist Maj. Nidal Hasan, allegedly opened fire at the Soldier Readiness Center, killing thirteen people and wounding forty-three others. According to some accounts, there is evidence that Maj. Hasan had contact with radical Islamist elements before his shooting spree.[28] In addition to this alleged cyber-use evidence, Maj. Hasan exhibited behavioral and sociological indicators that could have systematically alerted his associates and superiors to the high risk he posed. Thus, for example, the U.S. army Fort Hood Internal Review Team Final Report provides specific recommendations for pre- and post-deployment behavioral screening to improve communication among patients, providers, and commanders;[29] for health-care providers to provide information to commanders relating to indicators of possible violence; and for training of "... Soldiers to identify and report Soldiers that exhibit indicators of potential violence and/or potential terrorist behavior..." The technical approach to threat assessment that we have described in this article would certainly apply to this type of case, as well as to data leakage/exfiltration and espionage scenarios.

A second point concerning the generality of the approach is that the reasoning system developed here will address a broad class of decision analytic challenges facing the intelligence and counterintelligence communities today. Given the domain-independent structure of the CHAMPION reasoning system, it is straightforward (but not trivial) to apply this technology to problems in counterterrorism, weapons nonproliferation, and related threat and warning analysis functions. The technical implementation underlying the reasoning framework is closely aligned with sense-making approaches to decision-support systems that have been promoted by the cognitive systems engineering research community, especially the general framework that Gary Klein and collaborators have described as "Recognition-Primed Decision Making" (RPDM).[30] Indeed, we suggest that the approach described in the present article offers one method of implementing an operational version of a RPDM model.

## Conclusion

The insider threat, especially espionage and data leakage involving computer networks, is among the most pressing cyber-security challenges that threaten government and industry information infrastructures. Unfortunately, no single intrusion detection or threat assessment technique in wide use today gives a complete picture of the insider threat problem. A predictive modeling approach to insider threat mitigation was described that aims to incorporate a diverse set of data sources that not only address

the cyber domain but also the psychological/motivational factors that may underlie malicious insider exploits. This comprehensive threat assessment framework promises to automate the detection of high-risk, concerning behaviors ("precursors" or "triggers") on which to focus the attention and inform the analysis of cyber-security personnel, who would otherwise be required to analyze and correlate an overwhelming amount of data. Incorporating psychosocial data along with cyber data into the analysis offers an additional dimension upon which to assess potential threats within a comprehensive, integrated threat analysis framework.

Current practice tends to be reactive, as it focuses on detecting malicious acts after they occur, with the aim of identifying and disciplining the perpetrator. In addition, the cyber-security/insider threat analysis process puts the greatest demand on the analyst for correlating multiple sources and patterns of data to recognize potential threats. We have developed a model-based belief propagation framework that uses psychosocial indicators as well as cyber indicators of potential abuse of network resources to identify and proactively respond to possible malicious exploits. Some indicators may be observed directly, while others are inferred or derived from observed data.

Defining triggers in terms of observable cyber and psychosocial indicators and higher-level aggregated patterns of these behaviors is a major challenge, but also a critical ingredient of a predictive methodology. The incorporation of psychosocial indicators is a serious matter that must be conducted with care to protect individual privacy. We have argued elsewhere for the need to apply ethical standards in limiting the type of information monitored and the access to such information by security analysts,[31] while advocating that this is an attainable goal that both protects individual privacy and enables the organization to protect its material, intellectual property, and personnel assets. Prerequisites for a valid deployment of this approach are effective training for management and HR personnel in recognizing and reporting behavioral precursors, establishment of policies and mechanisms for addressing insider threats through formation of interdisciplinary teams (representing management, HR, security, and counterintelligence perspectives), and proper safeguards and protection of the information from improper use or leakage. An informed and enlightened organization requires that management and HR staff be equipped with tools to maintain awareness of worker satisfaction and well-being. However, these tools cannot overstep ethical and privacy boundaries. This allows the organization to respond thoughtfully, effectively, and proactively to situations that, if unaddressed, may otherwise increase the risk of insider attacks. Further research and technology development along the lines described in this paper, as well as discussion

of social and ethical issues in employee monitoring, should remain among the highest priorities in addressing the insider threat.

## About the Authors

Dr. Frank L. Greitzer is a Chief Scientist at the Pacific Northwest National Laboratory (PNNL), where he conducts R&D in human decision-making for diverse problem domains. At PNNL Dr. Greitzer leads the cognitive informatics R&D focus area, which addresses human factors and social/behavioral science challenges through modeling and advanced engineering/computing approaches. This research focuses on the intelligence domain, including human behavior modeling with application to identifying/predicting malicious insider cyber activities, modeling socio-cultural factors as predictors of terrorist activities, and human information interaction concepts for enhancing intelligence analysis decision-making. Dr. Greitzer's research interests also include evaluation methods and metrics for assessing effectiveness of decision aids, analysis methods, and displays.

Ryan Hohimer is a Senior Research Scientist at PNNL. His research interests include knowledge representation and reasoning, probabilistic reasoning, semantic computing, cognitive modeling, image analysis, data management, and data acquisition and analysis. He is currently serving as Principal Investigator of a Laboratory-directed Research and Development project that has designed and developed the CHAMPION reasoner.

## Acknowledgments

The authors wish to thank Dr. Deborah A. Frincke, Lead for the Information and Infrastructure Integrity Initiative at the Pacific Northwest National Laboratory (PNNL). We also thank Christine F. Noonan for support in preparing this manuscript. This work was supported by the Information and Infrastructure Integrity Initiative of the Pacific Northwest National Laboratory. The Pacific Northwest National Laboratory is operated by Battelle for the U.S. Department of Energy under Contract DE-AC06-76RL01830. PNNL Information Release Number PNNL-SA-78381.

## References


1 G. Adams, "Private Memo Exposes US Fears over Wikileaks," *The Independent*, London, 2011, available at: *http://tinyurl.com/5w532h6 (www.independent.co.uk/news/world/americas/private-memo-exposes-us-fears-over-wikileaks-2177041.html).*

2 DoD Office of the Inspector General, *DoD Management of Information Assurance Efforts to Protect Automated Information Systems* (Washington, D.C.: U.S. Dept. of Defense, 1997).

3 U.S. Secret Service, Software Engineering Institute, CERT Program at Carnegie Mellon University and Deloitte, "2010 CyberSecurity Watch Survey—Survey Resuts," *CSO Magazine*, 2010, available at: *http://www.sei.cmu.edu/newsitems/cyber_sec_watch_2010_release.cfm*.

4 Verizon and U.S. Secret Service, "2010 Data Breach Investigations Report," 2010, available at: *http://tinyurl.com/26cqfj2 (www.verizonbusiness.com/resources/reports/rp_2010-data-breach-report_en_xg.pdf).*

5 F.L. Greitzer, T. Carroll, J.D. Fluckiger, L.J. Kangas, C.F. Noonan, and P.R. Paulson, *Insider Threat Modeling for Misuse Detection and Prevention (OUO)* (Richland, WA: Pacific Northwest National Laboratory, 2010).

6 F.L. Greitzer and D.A. Frincke, "Combining Traditional Cyber Security Audit Data with Psychosocial Data: Towards Predictive Modeling for Insider Threat," in: Christian W. Probst, Jeffrey Hunker, Dieter Gollmann, and Matt Bishop eds., *Insider Threats in Cyber Security* (New York: Springer, 2010), 85–114.

7 M.R. Randazzo, M. Keeney, E. Kowalski, D. Cappelli, and A.P. Moore, "Insider Threat Study: Illicit Cyber Activity in the Banking and Finance Sector" (Pittsburgh, PA: Carnegie Mellon University, 2005), available at: *http://www.sei.cmu.edu/reports/04tr021.pdf*.

8 E. Cole and S. Ring, *Insider Threat: Protecting the Enterprise from Sabotage, Spying, and Theft* (Rockland, MA: Syngress Publishing, 2006), 29.

9 "Internal review: The insider threat risk," *SC Magazine*, 2011, available at: *http://tinyurl.com/6xx6dyc (inform.com/government-and-politics/internal-review-insider-threat-risk-4737197a).*

10 G. Lawton, "New technology prevents data leakage," *IEEE Computer* 41 (2008): 14–7.

11 H.W.J. Rittel and M.M. Webber, "Dilemmas in a General Theory of Planning," *Policy Sciences* 4 (1973): 155–69, and H.W.J. Rittel and M.M. Webber, "Planning Problems Are Wicked Problems," in: N. Cross ed., *Developments in Design Methodology* (New York: Wiley, 1984), 135–44.

12 G. Klein, *Streetlights and Shadows: Searching for the Keys to Adaptive Decision Making* (Cambridge, MA: MIT Press, 2009).

13 F.E. Emery and E.L. Trist, *Towards a Social Ecology: Contextual Appreciation of the Future in the Present* (London: Plenum, 1972).

14 D.R. Band, D. Cappelli, L.F. Fischer, A.P. Moore, E.D. Shaw, and R.F. Trzeciak, *Comparing Insider IT Sabotage and Espionage: A Model-based Analysis* (Carnegie-Mellon University: Software Engineering Institute. CERT Coordination Center, 2006), available at: *http://www.cert.org/archive/pdf/06tr026.pdf*.

15 Terry M. Gudaitis, "The Missing Link in Information Security: Three Dimensional Profiling," *CyberPsychology & Behavior* 1 (1998): 321–40.

16 E. Eugene Schultz, "A framework for understanding and predicting insider attacks," *Computers & Security* 21 (2002): 526–31.

17 C. Langin and S. Rahimi, "Soft Computing in Intrusion Detection: the State of the Art," *Journal of Ambient Intelligence and Humanized Computing* 1 (2010): 133–45.

18 K. Ilgun, R.A. Kemmerer, and P.A. Porras, "State Transition Analysis: A Rule-Based Intrusion Detection Approach," *IEEE Transactions on Software Engineering* 21 (1995): 181–99.

19 D.E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning* (Reading, MS: Addison-Wesley, 1989).

20 F.L. Greitzer, L.J. Kangas, C.F. Noonan, and A.C. Dalton, *Identifying At-Risk Employees: A Behavioral Model for Predicting Insider Threats* (Richland, WA: Pacific Northwest National Laboratory, PNNL-19665, 2010); F.L. Greitzer and D.A. Frincke, "Combining Traditional Cyber Security Audit Data with Psychosocial Data: Towards Predictive Modeling for Insider Threat."

21 M. Dean, G. Schreiber, S. Bechhofer, F. van Harmelen, J. Hendler, I. Horrocks, D.L. McGuinness, P.F. Patel-Schneider, and L.A. Stein, "OWL Web Ontology Language Reference, W3C Recommendation," 2004, available at: *http://www.w3.org/TR/owl-ref/*.

22 J.A.R. Garcia, "jCOLIBRI: A multi-level platform for building and generating CBR systems," *Ph.D. Dissertation, Department of Software Engineering and Artificial Intelligence* (Madrid: Facultad de Informática Universidad Complutense de Madrid, 2008).

23 J.L. Kolodner, "An Introduction to Case-based Reasoning," *Artificial Intelligence Review* 6 (1992): 3–34; S. Bogaerts and D. Leake, "IUCBRF: A Framework for Rapid and Modular Case-based Reasoning System Development" (Bloomington, IN: Indiana University, 2005).

24 A. Aamodt and E. Plaza, "Case-based Reasoning: Foundational Issues, Methodological Variations, and System Approaches," *AI Communications* 7 (1994): 39–59.

25 I. Horrocks, P.F. Patel-Schneider, H. Boley, S. Tabet, B. Grosof, and M. Dean, "SWRL: A Semantic Web Rule Language Combining OWL and RuleML, W3C Member Submission," 2004, available at: *http://www.w3.org/Submission/SWRL/*.

26 D. Nardi and R.J. Brachman, "An Introduction to Description Logics," in: F. Baader, D. Calvaneses, D.L. McGuinness, D. Nardi; and P.F. Patel-Schneider, eds., *The Description Logic Handbook* (Cambridge: Cambridge University Press, 2003), 5–44.

27 F.L. Greitzer, T. Carroll, J.D. Fluckiger, L.J. Kangas, C.F. Noonan, and P.R. Paulson, "Insider Threat Modeling for Misuse Detection and Prevention (OUO);" F.L. Greitzer and D.A. Frincke, "Combining Traditional Cyber Security Audit Data with Psychosocial Data: Towards Predictive Modeling for Insider Threat."

28 E. Perez and K. Johnson, "Hasan, Radical Cleric Had Contact," *The Wall Street Journal*, 2009, available at: *http://online.wsj.com/article/SB125778227582138829.html*.

29 U.S. Army, "Fort Hood Army Internal Review Team: Final Report," *U.S. Department of the Army*, 2010, available at: *http://tinyurl.com/5uug3rn (www.globalsecurity.org/military/library/report/2010/ft-hood_airt_final-report.htm)*.

30 G.A. Klein, "A Recognition Primed Decision (RPD) Model of Rapid Decision Making," in: G.A. Klein, J. Orasanu, R. Calderwood, and C.E. Zsambok, eds., *Decision Making in Action: Models and Methods* (Norwood, NJ: Ablex, 1993), 138–47.

31 F.L. Greitzer, D.A. Frincke, and M.M. Zabriskie, "Social/Ethical Issues in Predicitve Insider Threat Monitoring," in: M.J. Dark, ed., *Information Assurance and Security Ethics in Complex Systems: Interdisciplinary Perspectives* (IGI Global, 2010), 132–61.

Journal of Strategic Security