Training organizational supervisors to detect and prevent cyber insider threats: two approaches

Dee H. Andrews *, 1, 4, Jared Freeman 2, Terence S. Andre 3, John Feeney 2, Alan Carlin 2, Cali M. Fidopiastis 3, Patricia Fitzgerald 4

1 Army Research Institute, 425 E. Melody Lane, Gilbert, Arizona, USA, 85234
2 Aptima, Inc.
3 Tier1 Performance Solutions, Inc.
4 Formerly with the Air Force Research Laboratory

Abstract

Cyber insider threat is intentional theft from, or sabotage of, a cyber system by someone within the organization. This article explores the use of advanced cognitive and instructional principles to accelerate learning in organizational supervisors to mitigate the cyber threat. It examines the potential advantage of using serious games to engage supervisors. It also posits two systematic instructional approaches for this training challenge – optimal path modelling and a competency-based approach. The paper concludes by discussing challenges of evaluating training for seldom occurring real world phenomena, like detecting a cyber-insider threat.

Keywords: accelerated learning, cognitive principles, cyber insider threat, game-based instruction.

Received on 26 March 2012; accepted on 26 March 2013, published on 03 May 2013

Copyright © 2013 Andrews et al., licensed to ICST. This is an open access article distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/3.0/), which permits unlimited use, distribution and reproduction in any medium so long as the original work is properly cited.

doi: 10.4108/trans.sesa.01-06.2013.e4

1. Introduction

Cyber theft and sabotage are a rising threat in our increasingly wired world. Vast sums of money are spent by organizations and individuals who do not want their sensitive cyber information pursued and stolen. Yet, cyber insider damage and theft happens on a regular basis. Software solutions that purport to solve the problem are available, yet the problem persists. One major reason is because small minorities of individuals inside organizations choose to betray the trust they have been given and commit insider threat acts.

Cyber insider threat accounts for a large percentage of total cyber threats. For example, "The 2011 CyberSecurity Watch Survey uncovered that more attacks (58%) are caused by outsiders (those without authorized access to network systems and data) versus 21% of attacks caused by insiders (employees or contractors with authorized access) and 21% from an unknown source; 33% view the insider attacks to be more costly, compared to 51% in 2010. Insider attacks are becoming more sophisticated, with a growing number of insiders (22%) using rootkits or hacker tools compared to 9% in 2010, as these tools are increasingly automated and readily available." (Software Engineering Institute, Carnegie Mellon University, 2011, p. 1).

* Corresponding author. Email: dee.h.andrews@us.army.mil
In a now infamous case, the WikiLeaks website is alleged to have gained access to a very large number of classified files via an insider. (Greenemeier & Choi, 2010).

Currently, few supervisors in the workplace receive any training about how they can detect and help prevent cyber insider threats, including information on the types, impacts, or prevention of those threats. Supervisors are busy people who don’t have much time to be trained on the cyber threat or about how to mitigate the threat potential. They often are not trained in information technology, so they must rely on information technology specialists in detecting suspicious activity via defensive software. However, supervisors can still play a vital role in mitigating the threat by observing employee attitudes and behaviour, provided they know what to look for. From behavioural research cited in this paper can be constructed a list of concepts, principles and techniques that can be used by supervisors as they attempt to mitigate insider threat. Key questions are: How can we best train supervisors? Can we accelerate learning of these insider threat detection skills?

In this article we describe competencies for detecting and preventing cyber insider threats, and discuss key cognitive principles that can be used in developing and delivering training. We also discuss accelerated learning strategies for training busy supervisors.

A variety of factors may hinder a front line supervisor from detecting a cyber-insider threat. Supervisors are busy with their day-to-day supervisory activities in the workplace. Their attention is focused on the technical tasks and the management of their charges. In addition, most managers would rather not suspect their employees of malfeasance. It is much better to have an atmosphere of trust; distrust can ruin any workplace environment. However, even given those constraints, front line supervisors are in good positions to judge employee attitudes and behaviours. Perceptive advisors can often tell when something is bothering an employee, and with the proper training can work with security, information technology and human resource professionals to detect behavioural problems that could lead to insider threat activities, of both the cyber and non-cyber varieties. Hopefully, early detection of these threats will not only mitigate their possibility and impact, but also lead to help for the employee. Timely, appropriate and effective training for supervisors can play an important role in protecting cyber systems and vital information contained therein.

2. Accelerated learning for training

Due to the critical nature of the task of mitigating insider threat, and due to the time pressures on organizational supervisors it is vital that any training that supervisors receive be shaped so that it can accelerate learning. Accelerated learning can be very helpful in a variety of learning settings. For example, one critical future need in the military is to dramatically and effectively accelerate the transition from novice to expert in the cyber domain, which would have a significant impact on military personnel performance and overall cyber mission effectiveness (Department of Defense, 2009). Lozanov (1978) described the construct of accelerated learning as a mechanism: an assistive paraconscious mental activity to automate and use memory, brain, and intellectual reserves of people more effectively. Others define the construct by its effects: as a change in the speed with which a learner can process and retain to-be-learned material (Rose & Nicholl, 1997; Russell, 1999; Landale, 2004; Lawlor & Handley, 1996). Both definitions address the learner’s ability to rapidly encode information during learning and retrieve it during task execution. Hoffman, Ward, DiBello, Feltovich, Fiore, and Andrews (in press) provide three definitions of accelerated learning from different perspectives: The first is “rapidized training”—the idea of training individuals to achieve some minimal level of proficiency at a rate faster than usual. Second, accelerated learning also refers to the idea of getting individuals to achieve high levels of proficiency at a rate faster than ordinary. The most succinct way of saying this is to ask: “Can we turn an apprentice into an expert in less than ten years?” Third, accelerated learning also refers to the idea of making learning more immune to decay. In other words, once trained to a high level of proficiency, how can one at least stabilize that level of skill?

2.1 Serious games for accelerated learning and learner engagement

One instructional approach to accelerated learning is to encapsulate the learning in an immersive environment, providing the learners with some simulated real-world examples of insider threat behaviour. For example, serious games may be able to increase learner interest in the skills and knowledge necessary to mitigate cyber insider threat.

Game-based learning environments, also known as immersive learning simulations or serious games, provide a possible approach to train knowledge, skills, and abilities (KSAs) that lead to better learning and transfer of training to the operational environments. The eLearning Guild (2008) defines serious games as, “an optimized blend of simulation, game element, and pedagogy that leads to the learner being motivated by, and immersed into, the purpose and goals of a learning interaction.” Serious game developers use meaningful contextualization, and optimized experience, to successfully integrate the engagement of well-designed games with serious learning goals. Game-based learning environments could be designed to increase student engagement (Bergeron, 2005). However, measures of learner engagement are typically subjective. Psychophysiological measurements such as EEG and eye-tracking sensors have led to significant advances in scientists’ understanding of focusing attention during training (Fox, 2008; Rock & Schwartz, 2006; Rock & Schwartz, 2007). Through accelerated learning approaches and innovative ways to measure learner engagement, training developers can establish new guidelines for creating game or simulation-based training (Fidopiastis & Nicholson, 2008).
The cyber insider threat challenge described above has two aspects, operational and scientific. The operational objective is to improve the ability of front-line managers to detect and act on potential insider threats. To successfully mitigate insider threat the supervisor must develop three skills: assessing the threat level, attributing an assessment to learned threat cues, and acting to manage the threat.

The scientific challenge is to create and validate a generalizable solution to the challenge of training in ill-defined domains such as the insider threat. This challenge can be addressed by defining an instructional strategy and a technology to implement it. A viable strategy is deliberate practice, in which the student receives didactic instruction concerning aspects of expert knowledge and practice, performs the task while attending to these elements, receives remedial feedback, and continues practice and feedback until some level of expertise is achieved. This strategy helps the learner to efficiently encode knowledge of key domain concepts and solution procedures and, by extension, facilitates near-automatic retrieval of knowledge and execution of skills. Deliberate practice has produced strong learning outcomes across a wide range of cognitive and psychomotor tasks (Ericsson, Krampe, & Tesch-Romer, 1993). The technology that implements this strategy enables the instructional designer to specify the structure of training domains, organize training activities within that structure, dynamically optimize the sequential delivery of activities as a function of measured learner performance, and provide feedback based on that measured performance.

One way to apply a serious game is to give the trainee a brief didactic instruction concerning the importance of the insider threats and cues to detecting the threat, then have the student solve brief mystery problems delivered in a dynamically optimized sequence. Each scenario could consist of cues that are delivered in email, voicemail, and documents. The student investigates these materials briefly (e.g., for 5 minutes), then responds to a test in which s/he assesses the threat, explains (or attributes) that assessment to common causes (below), and takes recommend actions (e.g., meet with the employee, engage human resources, or notify security). The serious game then would provide feedback and direct the student to scenarios that would optimize their learning and variety of experience over time. The training game should also provide access to organization-specific reference materials (e.g., security policies and procedures).

Frank Greitzer and colleagues have published three reports that describe psychosocial cues to insider threats. (Greitzer, Kangas, Noonan, Dalton., & Hohimer, 2012; Greitzer, Frincke, & Zabriskie, 2011; Greitzer & Frincke, 2010). The research they describe concerns environmental factors that influence counterproductive workplace behaviours (c.f., Tripp & Bies, 2009; Fox & Spector, 2005; Katz & Kahn, 1978), and common security policies. An accelerated training program for supervisors should seek to develop student knowledge of these three domains such that they can assess the presence (or absence) of insider threats accurately; attribute those assessments correctly to psychosocial, environmental, and security cues; and select appropriate actions. Domain expertise consists, more specifically, of knowledge of 28 types of cues:

- Twelve psychosocial cues reliably recognized by security and human resources professionals, cues such as disgruntlement, disregard for authority, disengagement, and stress;
- Ten environmental cues concerning workplace stressors (e.g., recent downsizing, lack of recognition for performance), managerial practices (e.g., poor communication with staff), and organizational governance (e.g., lack of a functioning organizational justice system, lack of response to poor performance or behaviours);
- Six security cues concerning the existence, understanding, and enforcement of security policy.

Game based training scenarios should be systematically designed to present cues to none or several of these 28 competencies. The scenarios should sample the combinatorial space of competencies well within the constraints imposed by the brevity of scenarios and cost of scenario production, review, and revision.

3.2 Training objectives from the case literature

An alternative strategy is to define training objectives from a body of case studies in insider cyber threat. Table 1 presents objectives based on case studies published by Carnegie Mellon’s Computer Emergency Response Team (CERT), (Cappelli, Moore, Shimeall & Trzcieniak, 2006).

3. Training objectives

A significant challenge in designing such training would lie in defining the competencies students must learn. Two general strategies that appeal to the authors are: 1. mine the scant research literature for skill requirements, and 2. leverage the case literature from insider threat incidents.

3.1 Training objectives from the research literature
Table 1. Insider threat competency model.

<table>
<thead>
<tr>
<th>Competency</th>
<th>Source</th>
</tr>
</thead>
</table>
| Be proficient in general security awareness | CERT Observation #5  
Insiders created or used access paths unknown to management to set up their attack and conceal their identity or actions. The majority attacked after termination. |
| Be able to recognize employee predispositions | CERT Observation #1  
Most insiders had personal predispositions that contributed to their risk of committing malicious acts. |
| Be able to recognize employee behavioural patterns, general appearance, and physical health | CERT Observation #4  
Behavioural precursors were often observable in insider IT sabotage cases but ignored by the organization. |
| Be able to manage employee expectations | CERT Observation #2  
Most insiders’ disgruntlement is due to unmet expectations. |
| Be proficient in detecting an insider threat | CERT Observation #6  
In many cases, organizations failed to detect technical precursors.  
Awareness of technical precursors and actions performed from within the organization that point to signs of insider threat: |
| Be able to detect individuals who are high risk and high liability to security during the hiring process | Interview with Air Force Computer Crime Investigator Subject Matter Expert  
Consider insider threat when hiring someone. |

From this initial model, we have identified over 100 behaviours that have been associated with the successful detection and mitigation of insider threats to cyber security. From the original list of competencies and the list of behaviours identified in the literature, the authors propose a new model for review. This new model groups behaviour statements that are then collapsed into five categories reflective of competencies for the domain investigated. This model has undergone a Q-sort analysis and has been revised based on critiques by cyber security experts in the military and at CERT. As a result of this extensive review and amalgamation of expert opinions, the authors have updated the original list of competency statements (left column of Table 2) to a revised set (right column of Table 2).

Table 1: Comparison of original and revised competency models for insider threat detection

<table>
<thead>
<tr>
<th>Original List of Competencies</th>
<th>Revised List of Competencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Be proficient in general security awareness</td>
<td>Proficiency in general cyber security awareness</td>
</tr>
<tr>
<td>Be able to recognize employee predispositions</td>
<td>Ability to detect and reject high risk applicants prior to hire/placement</td>
</tr>
<tr>
<td>Be able to recognize employee behavioural patterns, general appearance, and physical health</td>
<td>Proficiency in insider threat prevention or mitigation</td>
</tr>
<tr>
<td>Be able to manage employee expectations</td>
<td>Proficiency in detecting insider threats</td>
</tr>
<tr>
<td>Be proficient in detecting an insider threat</td>
<td>Proficiency in determining whether and/or how to respond to potential insider threat (or possible behavioural or technical predictor thereof)</td>
</tr>
<tr>
<td>Be able to detect individuals who are high risk and high liability to security during the hiring process</td>
<td></td>
</tr>
</tbody>
</table>

4. Two alternative approaches to accelerating learning to mitigate insider threat

To ensure that a serious game based approach to training supervisors will be effective it will be necessary to adopt a training strategy that taps key cognitive principles in designing the instruction. Two possible approaches are presented here. One concentrates on defining a specific instructional architecture based on optimal path modelling, and the other focuses on a method called “content filtering”.

4.1 Dynamic selection of instructional events using optimal path modelling

The training should be built upon a generalizable instructional architecture. Its components would be an Instructional Strategy Model, which selects scenario parameters that optimally challenge and advance the trainee as a function of prior and expected performance; a Scenario Management Model that selects a scenario from a library on the basis of those parameters using a simple index of parameter values to scenarios; and a Game Environment, which would present didactic materials, scenarios, tests, and feedback based on the output of a performance measurement.
module. Figure 1 presents the instructional architecture for this approach.

![Instructional Architecture Diagram]

Figure 1. Instructional architecture.

The Instructional Strategy Model is central to this architecture and its use. To configure that model, the instructional designer would define the training domain, training scenarios, performance measures, and the desired balance of deep vs. broad expertise. To apply that model, we would use the model to generate a training policy that dynamically selects the scenarios that most rapidly advance a student towards expertise. The domain can be modelled as a Partially Observable Markov Decision Process (POMDP), (Sondik, 1971). This is one of a family of models used for optimal path planning. Here, it would be used (after Atkinson, 1972) to pilot the student through the space of training objectives given uncertain estimates of student knowledge state (e.g., of knowledge of psychosocial, environmental, and security factors) and the probabilistic effects of instructional actions (i.e., presentation of scenarios) on that knowledge. Levinchuk, Shebilske, and Freeman (2012) compared this dynamic planning method against a conventional strategy (static part-task hierarchical training) and found that the POMDP policy reliably accelerated learning and increased the adaptability of subjects to novel problems.

The training model, specifically, would represent the 28 training objectives (above), each at as many as four levels of salience across 76 scenarios. The levels would be high salience (i.e., cues obvious even to novices), moderate, low (i.e., cues detectable mainly by experts), or not present. Chronic and documented issues would constitute high salience cues; infrequent but unexplained issues would represent moderate salience; and infrequent or explained issues represent low salience. About three competencies would be presented in each scenario, with a minimum of one and a maximum of six competencies. The majority of scenarios would contain insider threat events (perhaps 80-90%); about 10-20% would have emergency (high level) threats; and about 5-10% would pose no threat. These data would be used were used to configure the Instructional Strategy Model and compute an optimal policy for dynamically selecting scenarios based on student performance.

### 4.2 Competency-based accelerated learning

A different strategy for accelerating learning for supervisors would use a dual approach, drawing upon findings in the literature and performance improvement expertise. This approach would include:

- Accelerating the learning pathway using a concept called “content filtering” to modify the learning pathway based on learner performance; and
- Accelerating the learning process itself through the application of research-based design principles found to support accelerated learning.

Content filtering, or sometimes called computer adaptive learning and assessment, describes the acceleration of instruction through the methodology of pretesting based on established learning objectives and then adapting the pathway based on results. This dual approach to accelerate the learning process would focus on foundational design elements applied within a delivery system for training supervisors to detect insider threats. These foundational design elements are:

(i) Scenario-based game engine. This game engine provides:

- Interactive scenarios that assess proficiency in each competency
- Practice in applying skills in a realistic environment
- After-Action Reviews to analyse performance

(ii) Learning pathways. These pathways are adapted by:

- Identifying all competencies and learning objectives for the training
- Using a pre-test to analyse learner proficiency of competencies and learning objectives prior to training
- Adapting the learning pathway by selecting those lessons applicable to the student needs

(iii) Instructional design best practices to support accelerated learning and engagement. These include:

- Learner Centered Focus (gain and keep learner attention; address multiple learning styles; encourage learning; allow for practice)
- Emotional Engagement (multiple perspectives/stories; emotion included in the experience; conversational writing style)
Higher Level Thinking (application; analysis; evaluation; reflection)

These design elements would be integrated into a scenario-based game environment to accelerate both the learning pathway and the overall learning process. Figure 2 shows how these design elements might be integrated to accelerate learning.

5. Evaluating training for seldom occurring tasks

We conclude by discussing a problem that is a challenge for training tasks that rarely occur in the on-the-job environment. Namely, how can we know if the training is effective if those who are trained rarely, if ever, are called upon to perform that task, in this case, detecting a cyber-insider threat. Detecting cyber threats in a busy workplace requires not just good training in spotting important insider cues, but it also requires a certain amount of vigilance on the part of the supervisor. Yet, non-cyber information technology supervisors are typically focused on issues unrelated to insider threats so even good training may not be enough for the detection task.

Kirkpatrick (1996) presented four levels of training evaluation:

- **Level 1: Reaction** - How well did the learners like the learning process?
- **Level 2: Learning** - What did they learn? (the extent to which the learners gain knowledge and skills)
- **Level 3: Behaviour** - (What changes in job performance resulted from the learning process? (capability to perform the newly learned skills while on the job)
- **Level 4: Results** - What are the tangible results of the learning process in terms of reduced cost, improved quality, increased production, efficiency, etc.?

Were training produced by either of the two methods described earlier in this paper, it would be relatively easy to evaluate the training at Levels 1 and 2 as soon as the trainee finished the training. However, evaluating the training’s effect after the trainee returns to the job is difficult for Levels 3 and 4. Insider threat cues are often subtle and they may not be obvious to a supervisor who is focused on other matters. Just because a supervisor does not detect an insider threat does that mean the training was not effective? Safety training presents a similar challenge, “Just because an accident happened, does that mean the safety training was not effective?” Cyber insider damage, like accidents, seldom occurs but both can have devastating consequences.

For such task domains, training evaluation must be done over a long period. To evaluate both behaviours and results for insider threat detection the evaluator would need to have at least a five year time frame, and a data collection system that would follow many dozens, perhaps hundreds of training supervisors, to collect enough data to make an effectiveness decision. The problem is compounded by the fact that many organizations would be hesitant to allow an evaluator from outside the organization to collect such data. Naturally organizations would prefer that insider threats not be revealed to customers or the public since doing so might erode confidence in the organization. Evaluators would obviously have to go to great lengths to assure the organization of the confidentiality of such information.

One intriguing option is to plant a confederate insider threat within a team, perhaps as a temporary worker or maintenance technician, and to assess the responsiveness of supervisors and staff to the threat. This strategy is currently used to test the responsiveness of staff within some DoD contractors to phishing designed to elicit export-controlled information about defense technologies. It is not, to our knowledge, used to test detection of insider threats, nor is the impact of such ploys on permanent staff and their operations known.

Despite the challenges in evaluating the effectiveness of insider threat training, the authors urge trainers to make the effort. If the training approach, be it the two outlined in this paper, or some other, is not evaluated the instructional field will have great difficulty in improving this vital area of training.

Acknowledgements

This work was funded by the Office of the Assistant Secretary of Defense (Research and Engineering). The opinions expressed here are the authors’ and do not necessarily reflect the policy of the Office of the Secretary of Defense.
References


