Decision Support Procedure in the Insider Threat Domain

John P. Murphy, Vincent H. Berk, Ian Gregorio-de Souza
Thayer School of Engineering
Dartmouth College
Hanover, NH
Email: firstname.lastname@dartmouth.edu

***Abstract*—Effective mitigation of the Insider Threat in complex organizations is not simply a matter of 'fire-and-forget'. Thorough routines are required to minimize the chances of malicious insiders going undetected. While detecting policy violations and signatures of known-bad behavior are essential to a broader threat mitigation strategy, it is clear that behavior-based measurements, including anomaly detection and social network analysis, will be crucial to detecting technically savvy malicious users with legitimate network and data access. Due to the large number of potentially malicious behaviors users may display, the main thrust of detection falls in the hands of an analyst capable of correlating these behaviors. Based on our BANDIT system, we offer a 10-step analyst program, which offers a common-sense approach to limiting the damage a malicious trusted user can achieve.**

***Keywords*-behavioral anomaly detection, insider threat, risk mitigation**

I. Introduction

Among the unique features of the malicious insider problem is the difficulty in applying automated solutions. A potential insider may play any of a number of roles within an organization, and a number of potential malicious missions have been identified. Indeed, the term *insider threat* is not well-defined. This situation has resulted in a proliferation of signature- and anomaly-based methods for identifying some aspect of the malicious behavior. While detecting policy violations and signatures of known-bad behavior are essential to broader threat mitigation strategies, behavior-based measurements, including anomaly detection and social network analysis, are crucial to detecting technically savvy malicious users with legitimate network and data access.

Due to the sensitive and subtle nature of this threat, analyst-in-the-loop systems are generally agreed upon to be required, but the proliferation of signals threatens to overwhelm analysts. Moreover, the analyst's work only begins at the point where a user is identified as *"of-interest"*; at which point the analyst must identify and document the potentially threatening actions and establish context so that other departments (IT, Legal, HR) can decide and take action. Initiating a full investigation of a user is costly, time-consuming, and can reduce morale. Skilled analysts can effectively conduct an ad-hoc investigation through original logs, netflow records, and packet captures, but they are costly in terms of manpower and computing resources.

The BANDIT system provides a simple and intuitive organizational method for a large number of insider threat metrics. Although it is designed with the needs of anomaly detection methods in mind, including signature detectors and non-automated measures is straight-forward. It includes a database structure to facilitate the guided discovery process and allows the analyst to make full use of underlying data such that the discovery of a potential malicious insider leads the analyst to explore high-level scores down to raw data within the same framework. This saves analyst time, facilitates the collection of data to support security decisions, and ultimately streamlines the process of identifying and dealing with the most risky users.

Efficiency in the insider threat domain is no different than any other risk mitigation domain; the design and implementation of proper procedure is just as important as the tools deployed. Procedure allows the effective use of tools, proper vetting of false positives, and an understanding of the limitations by the operators.

This article is a companion to our HICSS publication[1] describing the internals of our BANDIT insider threat decision support system. We outline the steps and procedures necessary to make optimal use of the BANDIT system in larger enterprise environments.

II. Background

Approaches to the automated detection of malicious insiders have left the field somewhat fragmented. While most organizations share risk factors of sabotage and exfiltration of sensitive data to some degree, various sectors have different motivations for detecting these insiders: the banking industry focuses on internal fraud, technology companies on the theft of intellectual property, hospitals on patient confidentiality, and government on counter-intelligence. The sheer number of facets to the insider threat problem, combined with the number of potential missions has led to a proliferation of individual detectors, each geared to identify specific behaviors and patterns.

Despite the differences in emphasis, insider threat mitigation has three features in common:

1) Multiple data sources from host, network, and non-automated sources are used in order to get a multi-dimentional view that is harder to evade;

DOI 10.1109/SPW.2012.17

2) Analysis is centralized, so that individual hosts are trusted only with the collection of (some) data, with data stored and collected elsewhere for fusion and analysis; and
3) The human in the loop is emphasized both to make use of human cognition and judgment, and to guard against wrongly pursuing false or meaningless positives.

This has held true from early systems such as IDES[2] through to present day systems such as Intel's upcoming Pro-I software and MITRE's ELICIT [3][4]. The structure of the day-to-day efforts using these systems tends to be the same: a set of front-line analysts examine the results of alerts generated from automatic examination of a large-scale data collection regime. When those alerts and other results turn out to be of interest, the data in question is passed up the hierarchy for follow-up analysis.

Although efforts to baseline normal behavior at the host level date back to the early 1990s (IDES) for use in anomaly detection, signature-based methods have predominated in the security domain. Tools, such as Snort [5], are commonly used. These solutions are frequently deployed since focusing on known-bad behaviors allows them to catch low-hanging fruit which are intuitively easier to grasp, as each signature can be documented with explanation and procedure. These tools, however, often are set up as single-shot detectors and make it difficult to correlate multiple alerts from different sensors. For example, associating a suspicious log entry and a known-bad network packet sequence to the same user.

Systems, such as ELICIT, make use of scores from such detectors, plus simpler anomaly detection techniques. They often perform the important work of maintaining user identity consistency so that an analyst who sees multiple detector results can quickly determine which user is involved. As long as detectors are designed to clearly signal bad or risky behavior, then the volume of triggered detectors flags a misbehaving user. These systems are generally not provided as pre-packaged software solutions, however, and implementations are left to the deploying organization. This inconsistency has been the primary impediment to widespread success.

The deployment of large-scale anomaly detection that takes into account long-term behavior and social network analysis is hampered by the ability to effectively organize this information. Anomaly detection can produce false positives – a user's behavior is only anomalous because information such as meeting scheduling is not included – and a high meaningless true positive rate – that is, a user's behavior may be genuinely anomalous, but not in a dangerous way. One of the reasons for the prevalence of signature-based methods in this domain is because following up on positives is expensive: Expecting analysts to follow up on every flagged anomaly is impractical, and although promising, anomaly detection methods have not been widely used to date.

We argue that a flexible decision support system, such as BANDIT, combined with a rigid analyst process, can fuse the available data sources, and enable the cognitive abilities of the analyst to make significant headway in mitigating the threat of malicious insiders.

III. Procedure and Rationale

We split the process of Insider Threat analysis into three broad categories; strategy, tactical, and audit. The *Preparation* section sets the strategy for defense, sensing, and training. The *Daily Procedure and Reporting* section treats analysts' day-to-day tactical activities to detect and mitigate insider threats. The last section *Ongoing Audit*, reflects on the strategic and tactical processes, and how to improve them by making changes as necessary.

A. Preparation

1) Inventory your data sources and determine where the blind spots are: No detector can see everything. Capturing network traffic to the printer won't catch the user walking out with the document on a thumb drive. The keycard scanner at the front door won't catch a remote exploit on your network. You will necessarily have a range of sensing technologies on the network, including passive sensors such as log files. As the network and infrastructure is mapped out, and sensing technologies are inventoried, a clear picture of what can be monitored will develop. Charting this information creates a powerful training tool for the analysts, and a deeper understanding of the semantics of alerts.

2) Set your defenses: Armed with good insight into the network infrastructure, and its blind spots, you can strategically plan your defenses. Eliminating exfiltration and sabotage options that deployed sensors are unable to detect, will force malicious insiders into the open. This acts both as a deterrent as well as speeding up an eventual investigation.

3) Determine detectors and metrics that can be based on your sensors: Some sensors will signal undesired conditions directly (e.g. signature detection systems), however most sensors will require some additional processing (e.g. log files). A measurement of pages printed per user per day, will need a metric that translates the raw numbers into a range from 'normal' to 'suspicious'. These metrics can be simple thresholds, linear mappings, or possibly a complex behavioral profile. For instance, some people print more in their course of duties than others. It is the change in behavior that may signal an anomaly, not the specific behavior itself.

4) Organize detectors and metrics: The first and foremost benefit of the BANDIT system is the amortization of results from the various detectors into a comprehensive behavioral analysis score. By combining the user's typical features from many sensors, and comparing to the user's behavior in the past, as well as the user's behavior as part of his/her peers, a

score is created indicating how 'unusual' the user is behaving. By organizing the detectors and metrics accordingly into intuitive and recognizable categories, operator drill-down efficiency is improved.

This generally also quickly weeds out the majority of the false positives. For example, using a different printer may be caused by an exhausted toner cartridge in the usual printer. Other false positives might be due to holidays, fire drills, and assorted non-regular events that would cause a change in normal behavioral patterns. Having an intuitive way to get the the data that generated a high anomaly score will vastly improve investigation speeds, and analyst bandwidth.

5) Map detectors to expected behaviors: It is important to document the detectors and metrics according to specific behaviors they measure. It provides an intuitive severity to a detection, and guides investigative technique. It is important to realize that a high score by itself does not indicate a malicious insider. The BANDIT system is designed specifically to support operators in their investigation of unusual user activity. There can be many reasons for unusual behavior, and further analysis by the operator is required to establish user intent. A well documented detector-to-behavior mapping helps avoid needless low-yield investigations.

6) Training: The mappings and documentation created in the aforementioned procedures form the basis for analyst training. A key understanding of what each detector and metric provides, what can, and cannot be detected, and expected yield of each detector are the fundamentals of an efficient analyst investigation.

B. Daily Procedure and Reporting

The following section comprises our recommendation for using a tool such as BANDIT for guided discovery of users of interest – users who are of interest for other reasons (for example because they are about to or have recently left the organization, or have been flagged by Human Resources or other Security teams for review) can be investigated through such a tool, of course, but those users would be expected to have different priority evaluations.

One of the principles we apply here is that as much as possible should be done through a single user interface. Changing interfaces (such as to go from BANDIT to a separate log reader or packet capture) can introduce delay. Requiring the analyst to manually enter (and often retype) identifying information such as user ID, time frame, IP address, etc. creates opportunities for miscopying on both initial investigation and when reporting the investigation.

1) Prioritize users for investigation based on high-level scores: The job of the analyst is to determine which users are malicious insiders. Investigations take a lot of effort, and the analyst cannot expect to investigate each user all the time. A decision support tool like BANDIT can help prioritize analyst workload by providing a score indicating which users are behaving most anomalously (either with respect to their own history or to others), have violated the most policies (or the most important), and have tripped the most signature detectors (or the most severe). The software allows the analyst to view the users graphically, to identify the highest-scoring users and the users who are outliers. By graphing all the users with respect to their scores for personal change of behavior versus those users with respect to their group deviation, individuals who score highly on one axis but not another can be easily identified where they otherwise might not. The benefit to the analyst is an intuitive visual representation of the magnitude of change for each user.

2) Starting at the most interesting/risky users, determine which component scores and detectors caused the score to be high: Individual risk scores can be influenced by a lot of general odd behavior, or by a few key metrics. Seeing the list of violated policies or tripped signatures together can help an analyst grasp the nature of the underlying behavior, particularly when the time frame being examined is on the order of a day rather than a week. Anomaly detection metrics frequently lend themselves to their own visualizations – for example, temporal baselining techniques[6] can be graphed to show the user's ordinary daily behavior as contrasted by the anomalous activity.

After identifying users with risky behavior, the analyst drills down on the individual sensing technologies and metrics that caused the anomaly. Drilling down in this way can allow the analyst to quickly grasp the nature of the risky behavior. If proper inventory of the infrastructure was achieved in the *Preparation Stage*, ample documentation is available on each of the metrics, components and detectors, incorporating domain knowledge of the organization. While it is expected that each analyst should be familiar with all aspects of the system, some detections will be more common than others. In the realm of policy violations they may be extremely rare.

3) Examine the user's history with respect to high-level scores and component scores, familiarizing oneself with previous investigations (if any): User history provides vital context into ongoing anomaly or signature detection. Policy violations and anomalous behavior are frequently more interesting if they are ongoing, and so it is important that the analyst is familiar with historical data to determine trends and cyclic patterns. Graphing is generally the best way to achieve this, and it may be useful to provide additional mathematical and statistical analysis tools for this purpose.[7]

It is expected that the analyst will be referencing, where appropriate, the previous investigations into this user. In particular, where the user has a string of violations of the same policies or similar anomalies. By doing so, the analyst may take advantage of past time spent investigating, and can re-evaluate old notes to quickly rule out known false-positives.

By this stage, the analyst should have a sense of the

identified user's current behavior and any trends and historical patterns in the relevant behavior. If this user has a history of a specific type of policy violation, or has displayed increasingly erratic work hours over the last month, this is the stage at which the analyst should be cognizant of that.

4) Identify the user's peers: A significant part of determining the severity of a user's risk is to have a sense of how that user's behavior matches (or not) that of their peers. If, for example, everyone in a given section of the company starts committing the same policy violations, then the investigation needs to take place at a higher level, similarly, if there is evidence of a large-scale change in task. By the same token, if one user begins staying late and printing documents, such behavior will appear differently if the user is alone in the office. However, if the user has several peers also staying late, the behavior may quickly be considered benign. It is therefore key that operational information is available to the analyst that quickly identifies a user's work group or social group.

This information can come from a priori knowledge – groups of users who have been identified as belonging to the same group or working on the same task – or from social network analysis – groups of users identified because they frequently email each other, or have been identified based on resource selection as working on similar tasks. Discrepancies between groups generated in these two ways may be of interest to the analyst.

5) Examine the user's history with respect to their peers: Even if a user's behavior is significantly different from their peer-group, a check of historical comparison data can confirm that this is normal, or anomalous. The analyst must gather sufficient information about a user's peers with respect to the same metrics, such that a suitable determination can be made of the context in which the user of interest's actions occurred.

6) Determine whether to continue investigation into this user: At this point in the workflow, the analyst will have identified the reasons for the user's risk score with regard to anomalous behavior, what bad-behavior signatures were matched, and what policies may have been violated. Ultimately, the decision on next steps depend on organization-specific criteria and the detectors that flagged the bad behavior in the first place. Certain detectors may be flagged as prompting automatic investigations, such as access of a honey token file. Documentation of the detectors is key to this process.

Whether the investigation continues or not, documentation of the process up to this point is necessary so that future investigations have a reference and so that the decision can be revisited later without completely redoing all of the steps.

7) Examine sub-components and sub-detectors where appropriate, and go to underlying data according to personal expertise and indicated severity: If an investigation continues, at this point each anomalous metric and measurement may be studied. Where possible, the decision support system is used to perform detailed examination of the data underlying the flagged behavior, with as little extraneous data as possible. Other sources of information may be investigated as well. The analyst may look for log files, reports, or other out-of-band information that can conclude the investigation either way.

8) Document results of drill-down: Since historical investigative data is an important part of the analyst process, it is expected that the analyst keep meticulous track of the investigative process at all times. Organizational requirements will most often dictate the easiest process for doing so. The analyst should seek to provide a single document for later retrieval, complete with the time and author. Copious note-taking and annotation by the author, as well as room for other analysts to append information are important. We make no recommendations as to format, other than to suggest that investigative records should be easy to locate, and quickly searchable.

9) Make recommendation and pass along to superior for review: At this stage, the analyst should have examined the individual user in the context of their own past behavior and in the context of identified social and task peers. The exact criteria for whether to pass this user along for greater scrutiny is up to the organization, but will include risk factors such as the user's access to sensitive resources and permissions issues such as the sensitivity and privacy concerns of the data needed for further drill-down (e.g. full packet capture or HR files). Other practical factors, such as workload, may play a role in this step. Depending on the organization's preferences, this stage may be delayed until the analyst has investigated other users of interest, providing a ranking of the reports. The role of the software at this point is to provide easy review of the information collected so far, and the availability of any procedures and instructions as determined in the Preparation phase – for example, whether specific signatures require automatic elevation, or what data sources to append to the documentation if certain anomalies are detected.

10) Collate documentation and make available to other analysts: The final stage for the front-line analyst is to document the investigation thus far, regardless of conclusion. The facts of an investigation, even if not deemed interesting at the time, will be useful information in further investigations, particularly if annotations are available from different analysts. Corporate procedure is key, as information overload allows specific instances to slip through the investigative process. Clearly marked and easily searchable documentation of investigations are both a very sensitive, as well as very powerful asset in the insider threat mitigation strategy.

C. Ongoing Audit Process

Case studies indicate that while malicious insiders are capable of great damage, they are few and far between.

Furthermore, a number of potential indicators will have high false positive rates. Therefore, it is likely that the majority of investigations will result in a determination of "no threat", particularly when first deploying an insider threat mitigation protocol. Over time, however, analyst expertise and domain knowledge will develop. Making use of this expertise and knowledge to decrease workload and identify better practices requires an ongoing process of evaluation and audit. These audit steps should be performed regularly both to make best use of analyst expertise and new academic research.

1) Identify which detectors and metrics have resulted in the most and fewest alerts/investigations: The detectors and metrics that are likely to require the most attention in an audit process are those that trigger most often, and those that trigger the least often. Detectors that trigger too often will come to be seen as "safe" by analysts, whereas those that trigger rarely or not at all may be given too much weight when they do. But those detectors may just as easily be mis-configured, too generalized/specific, or otherwise not working as intended. Conversely, they may be working as intended, but not providing useful information. It is important to identify the reasons why detectors or metrics fall under this category because tuning the detector (weights, thresholds, etc) may provide the desired fix. Since effective sensor coverage is critical in this domain, eliminating detectors should be a last resort.

2) Discuss tuning, weighting, removing, updating, reclassifying with the team: The task of curating a suite of detectors and metrics involves a number of tasks. Organizations change over time in terms of personnel, resources, and task. Network traffic in particular tends to increase over time as applications become networked or migrate to the cloud, and as increased functionality and larger files become a part of web applications older detectors need to be reassessed periodically according to emerging needs.

When new technology such as new methods of communication and data transfer become available, they will need to be accounted for in existing metrics. New detectors and metrics will be needed for some. Continuous education of analysts on new technologies is a must. Understanding how they work, affect, and trigger detectors/metrics is key to relating alerts to benign or malicious activity. Analysts need to understand how new metrics relate to older ones with respect to importance and relevance. Ways in which these new detectors or metrics support or complement existing ones must also be explained.

3) Keep abreast of academic/professional discussions of existing and new metrics/detectors and new potential data sources; re-evaluate data blind spots and return to preparation stage where necessary: The insider threat community continues to develop new procedures and methods as the problem area is studied, making use of new insights into the problem domain, greater computational capability, new opportunities for malicious behavior, and new data sources. Best practices are expected to change over time as the insider threat problem evolves, and the malicious insiders themselves will change tactics according to their perception of the monitoring regime deployed against them. For these reasons, it is important to stay informed of new methods and approaches and for new evaluations of old approaches, and to reassess the organization's data collection, detector suite, and documentation in that light.

IV. Final Thoughts

The steps and procedures outlined here are designed to make maximal use of the time and expertise of an organization's trained analyst. We believe that they constitute both a best practices recommendation for creating an insider threat analysis group, as well as a blueprint for the design of the software and data collection required to effectively support such analysis.

A software tool that follows and supports these procedures, such as BANDIT, allows analysts to go deeper into the data without leaving the application, providing guidance so that front-line analysts can do more of the leg-work before passing on a lead to superiors or to HR or legal departments. This makes the greatest use of the time and energy of everyone involved, streamlines the processes to eliminate gaps and confusion, and ultimately will give organizations an improved ability to detect and mitigate the threat posed by malicious insiders.

References

[1] V. Berk, G. Cybenko, I. Gregorio-de Souza, and J. Murphy, "Managing malicious insider risk with bandit," in *Proceedings of HICSS 49*, 2012.

[2] T. Lunt, A. Tamaru, F. Gilham, R. Jagannathan, P. Neumann, and C. Jalali, "Ides: a progress report [intrusion-detection expert system]," in *Computer Security Applications Conference, 1990., Proceedings of the Sixth Annual*, dec 1990, pp. 273 – 285.

[3] D. Caputo, M. Maloof, and G. Stephens, "Detecting insider theft of trade secrets," *Security Privacy, IEEE*, vol. 7, no. 6, pp. 14 –21, nov.-dec. 2009.

[4] M. Maloof and G. Stephens, "Elicit : A system for detecting insiders who violate need-to-know," in *Recent Advances in Intrusion Detection*, ser. Lecture Notes in Computer Science, C. Kruegel, R. Lippmann, and A. Clark, Eds. Springer Berlin / Heidelberg, 2007, vol. 4637, pp. 146–166.

[5] "Snort - an open source network intrusion prevention and detection system." http://www.snort.org.

[6] J. Murphy, V. Berk, and I. Gregorio-deSouza, "Effectively identifying user profiles in network and host metrics." *Proceedings of the 2010 SPIE Conference on Defense, Security, and Sensing*, Apr 2010.

[7] D. Robinson, *PhD Thesis: Cyber-based Behavioral Modeling*. Thayer School of Engineering, Dartmouth College, 2010.