Winter 12-10-2016

Stopping Insiders before They Attack: Understanding Motivations and Drivers

Sanjay Goel
University at Albany, SUNY, goel@albany.edu

Kevin Williams
University at Albany, State University of New York, kwilliams@albany.edu

Stan Zavoyskiy
University at Albany, State University of New York, szavoyskiy@albany.edu

Recommended Citation

Goel, Sanjay; Williams, Kevin; and Zavoyskiy, Stan, "Stopping Insiders before They Attack: Understanding Motivations and Drivers" (2016). *WISP 2016 Proceedings*. 2.
http://aisel.aisnet.org/wisp2016/2

Stopping Insiders before They Attack: Understanding Motivations and Drivers

Research in Progress

Sanjay Goel

University at Albany, State University of New York Albany, NY, USA, {goel@albany.edu}

Kevin Williams

University at Albany, State University of New York Albany, NY, USA {kwilliams@albany.edu}

Stan Zavoyskiy

University at Albany, State University of New York, Albany, NY, USA
{szavoyskiy@albany.edu}

ABSTRACT

Insider attacks are able to evade traditional security controls because the perpetrators of the attack often have legitimate access to protected systems and data. Massive logging of user online activity data (e.g. file access or transfer, use of data storage devices, email records) is collected and analyzed to detect insider attacks (e.g. data theft, fraud, policy violation, etc.). Such techniques are fraught with drawbacks and limitations: 1) the proverbial "needle in a haystack problem," where very little useful information is found in massive data sets, especially where the incidence of malicious insider activities is very small compared to that of legitimate actors; 2) employee privacy issues may exist about the company monitoring employee behavior; and 3) these techniques are largely wanting in their accuracy, leading to notably high false positive rates. Perhaps the most salient limitation of these techniques is that the analyses are post-hoc, and by the time the activity is detected, the insider has already engaged in data theft or exfiltration, the impact of which may not be reversible. This paper discusses the concept of using probes for detection of threats, wherein user intentions to engage in insider attacks can be gauged by sending carefully designed probes that rouse malicious users into acting. In this research, we seek a broad understanding of the scope and relevance of such probes. There are various motivations for users to steal data, including financial gain, patriotic fervor, and disgruntlement with work. In the present experiment, we created simulated conditions to reflect common insider motivations by providing subjects with imagined scenarios, then asking them to take the perspective of insiders in those scenarios, and explicate their actions through a series of structured questions that mimic our probes. The results show the effect of different scenarios in motivating the users, and the effectiveness of different probes in eliciting their actions.

Key Words: insider threats, behavioral security, active probes, data theft

INTRODUCTION

Data theft has become a key security issue for organizations with danger of information leakage from both external hackers and insiders. While, protection of information is a challenge in itself, protecting information from insider threats is more challenging than from external threats since insiders have privileges that they can exploit to steal and misuse confidential information. Information theft is not a new problem. While the general behavioral traits and base motivations for insider information disclosure remain unchanged over time, the means for data theft have evolved considerably. The proliferation of electronic and media storage options, increasing network connectivity and emerging communication technologies have all increased the potential of data theft; prior methods of controlling exfiltration of data with physical controls and surveillance are no longer adequate; for instance, the small footprint of storage devices can make it easy for infiltrators to evade physical detection. Bradley Manning was able to exfiltrate all the information he stole by putting it on a compact disk disguised as a music CD. The surveillance and security strategies of intelligence agencies are based on existing vectors of attack; however, with emerging technologies these vectors are changing.

The motivation for data theft varies considerably; for instance motivation behind national espionage incidents is usually to steal military secrets for other countries and motivation for data theft from private firms is often personal gain. Attackers often rationalize their behavior as justified, perhaps to correct a perceived transgression ("the company is not treating me fairly") or out of a sense of entitlement ("I contributed heavily to this project so should be able to retain ownership"). Such rationalizations explain how people who may see themselves as moral and honest can engage in minor acts of data exfiltration (cf., Mazar, Omir and Ariely, 2008). While the individual acts are easy to self-justify, cumulatively, minor acts can result in significant damage to the firm.

In the current literature on insider threat behavior, the most commonly cited motives are personal financial gain, disgruntlement, and a sense of entitlement (Moore et al., 2011). Employees may be able to sell information to competitors, hackers, or criminal groups, and those who have taken a job with another company may steal data right before their departure, in the hope that access to the data will provide them leverage for a favorable standing within the new organization. Many inside attackers seem to act out of feelings of resentment over perceived injustices by their employer, either in terms of inequitable distribution of rewards (i.e., distributive injustice) or unjust treatment (i.e., procedural injustice; Willison, 2009). To them, stealing company secrets may be perceived as a way of restoring equity and fairness. Insider threats may also be motivated by feelings of social injustice, whereby workers seek to redress what they perceive as immoral acts on the part of their company. Finally, insider theft or attacks may also be motivated by employees' feelings of proprietorship over data created through their own work.

COUNTERING INSIDER MALICIOUS BEHAVIOR

Detecting and countering insider behavior has traditionally been achieved through forensics data analysis. Sensors are placed on the network and individual computers to collect data on user actions (e.g. file transfers, logins, USB usage), which are then analyzed for malicious behavior. Several forensic data analytic techniques have been used for insider behavior, such as Eldardiry et al. (2013), who propose a technique that uses sensor fusion techniques to analyze anomalies in user behavior. They flag anomalous patterns of behavior, such as when a user exhibits behavior that reflects activities of a group to which he/she does not belong, or when the user's behavior differs from that of peers in his/her own group. They use logging patterns, device and data access, email metadata, and search history in their data set and standard data mining techniques for the analysis. Legg et al. (2015) similarly use login, USB, email, web and file usage data to determine anomalies in the system; performing a series of analyses based on hourly and daily usage patterns. Shultz (2002) suggests a framework for insider behavior that utilizes regression/data mining on multiple data vectors including usage patterns, actions, meaningful errors (e.g. deleting log files), verbal behavior (e.g. hatred in emails, or hostility towards employer), and personality (e.g. introversion). A report from the defense company Raytheon on insider threats provides general guidelines for protection on a risk-based approach, including valuing assets, profiling individuals, investigating previous incidents, conducting surveillance on activities, and selectively analyzing data based on risk. In most of the research on insider threat based on data analytics, the data collected for analysis is very similar (i.e. file usage/access/transfer patterns, email meta data, search history etc.)

The fundamental challenge with most of these data analytic systems is that they rely on data analysis post-incident. The detection may take weeks or even months, although typically mitigation requires quickly controlling the damage and attempting to fix the breach to prevent future occurrences. Passive post-hoc analysis is not sufficient; identifying individuals who pose risks a priori and preventing insider theft from happening is certainly a more desirable approach. Developing user behavioral profiles and linking them to insider theft risk is one viable method that has been suggested in the literature. Symantec (2011) reports suggest the use of behavior profiles of employees in estimating and countering insider threat risk. There has been some work in the context of human behavior and propensity for insider theft. Moore et al. (2011) analyzed 48 insider theft incidents, concluding that individuals who fit into the category of an "ambitious leader" or an "entitled independent" are more likely to engage in data theft behaviors. Nurse et al. (2014) provide a broad framework for insider behavior research. They suggest identifying personality correlates related to a propensity for insider behavior. Research using the big five personality traits--openness, extraversion, conscientiousness, neuroticism, and agreeableness--has provided evidence that low conscientiousness and high neuroticism can predict an increased risk of malicious insider behavior (McCrae and Costa, 1990). Additionally, the dark triad of personality traits (Machiavellianism, narcissism, and psychopathy), as well as a lack of maturity, aggressiveness, poor social skills, personal integrity, and lack of self-esteem are also identified as traits that are likely to predict the greater likelihood of malicious insider behavior.

RESEARCH OBJECTIVES

Insider behavior is also driven by such intrinsic factors as fear and greed. A basic premise of our research is that, within a particular context, employees fall to their basal instincts and get motivated to commit data theft; and then wait for an opportune moment to realize their goals. Data theft culminates through the confluence of the motive, capability, and opportunity triumvirate. We expect that capability to exfiltrate data exists already and if not motivated hackers will be able to acquire it easily. Thus the goal of our research is to develop live probes that can serve as decoys, and can be used to see if a user is inclined to engage in data theft prior to actually exfiltrating data. Probes are designed to signal opportunity and thus stimulate theft-related activity. Employees who have a propensity to conduct insider theft will react to the probes and engage in data theft. If tools are put in place to measure insider behavior (e.g. file transfers and deletion, USB usage etc.) the probes will lead to identification of insider threats in the organization.

There are psychological attributes of humans that make them more vulnerable to insider behavior, and coincident with those psychological attributes, specific contextual situations that can motivate people to act as malicious insiders and steal data, such as: the downsizing of the company, a poor performance report, anger at the institution for unethical behavior or injustice, etc. Once users are motivated, they might seek an opportunity when they can steal data clandestinely without being observed, such as when computer systems are down. These motivations can be simulated by developing specific contexts that activate specific psychological traits of individuals.

In this study, we attempt to gain a broad understanding of this link between user motivation and actions based on specific cues (probes). We create various situational contexts within which such behavior is particularly likely to happen, and then introduce probes, and test their efficacy in spurring the behavior. The situational contexts serve to provide individuals with a motivational context to engage in malicious insider behavior, whereas the probes are designed to provide an opportunity, or instigating factor, for them to engage in data theft. In our studies, psychological motivations will be created through contexts, and probes will be developed to elicit insider reaction. We develop scenarios to simulate different contextual situations, and ask subjects to role play an insider and explicate their actions based on specific probes that are introduced. Each scenario has a well defined context followed by specific questions that elucidate how a user will respond in a contextual situation and how he/she will respond to the probes.

RESEARCH DESIGN

This research was designed to develop probes that will identify an individual who is at risk for insider data theft. The probes in most cases establish a user's propensity for data theft in a naturalistic setting by measuring his/her actions in specific situations. We also assume that differences in an individual's psychological attributes may play a role, such that the strategies and actions of insiders may vary based on these attributes. We devise several scenarios corresponding to typical employee motivations that drive insider behavior (as gleaned from the literature), and test the effectiveness of probes within each scenario.

Probes were manipulated via messages delivered to participant/insiders in the scenario. Five different email messages (probes) were created. Sent in the form of email messages, the probes included:

1. New security update coming in 3 days (signifying a brief window to act)
2. Security system down for maintenance (brief window of opportunity)
3. Random audit of computers announced (brief window to hide files or cover tracks)
4. File permissions changing (brief window to hide files or cover tracks)
5. Data vault/repository opened (opportunity to search for data)

Different motives for insider threat exist, and thus it was important to test the probes in different contexts. Different motives for insider threat were gathered from the literature and then narrowed down to five motives. The scenarios were written to capture the five different contexts, reflecting different potential precursors (motivations) to insider theft:

1. Financial Gain (new job offer)
2. Social Justice (anger at the organization for their policies and actions toward others)
3. Disgruntlement (perceived personal injustice; quit and find new job)
4. Patriotism/Loyalty (ex-patriate approached to steal data to advance a country's agenda)
5. Morality spy (ex-patriate upset by government's actions asked to exposing state secrets to representatives of foreign states)

In total, five scenarios and five probes were developed for the analysis, and surveys were created that will test each of the probes within each context. In the study, participants were asked to read the vignettes and to indicate how they think an insider will respond to specific probes in different contexts.

Each context was paired with each probe, creating a total of 25 scenarios representing unique context-probe combinations. It was not feasible or practical to recruit volunteers to complete all 25 scenarios in a within-subjects design. Nor was it feasible given our initial sample design a full factorial between-subjects design. Instead, we started by providing each subject with 3 scenarios representing unique context-probe combinations.

The scenarios in the surveys were divided into two parts. The first part of the scenario described the context: a description was given of a fictitious worker in a particular context who is considering engaging in insider theft.   For example, in the financial gain context, the worker is considering a job offer with a competitor, and to make his position stronger with his new employer he could take information on his current projects with him.

This section was followed by a series of questions asking about the likelihood of the worker engaging in seven different insider threat behaviors (i.e. taking notes on his research, talking to other researchers about their projects, looking for data on servers and data repositories, attempting to log into other computers to find important data; searching for files on servers and computers, taking pictures of product designs with a camera, downloading files to a personal USB). In the second part of the scenario, the fictitious employee is described as having decided to exfiltrate data, and then he/she receives an email message containing one of the five probes listed above. After reading the email message, participants were asked to rate the likelihood that the fictitious employee would engage in six threat-like behaviors (act immediately to secure data, search for files on servers or computers, attempt to log into different computers to find data, download data and files to a personal USB drive, use a camera to take pictures of product designs, and remove/delete files not related to one's own assigned projects). Each likelihood question was rated on a 5-point Likert scale ranging from 1 = Extremely Unlikely to 5 = Extremely Likely.

We collected data from cyber security experts at two security-related conferences: NYS Cyber Security Conference, and the Americas Conference on Information Systems. By including security experts, our goal was to ensure that the subjects would understand the concepts of insider data theft, as well as the ways that data can be exfiltrated. Online surveys were sent to employees of NYS Information Technology Services. Based on the feedback we received from the conferences, it was decided to reduce the number of scenarios from 3 to 2 for the online data collection, to reduce the burden on any individual subject. Our data collection strategy was not optimal from a design perspective, in that participants were exposed to only a small set of the possible experimental conditions. However, our goal was to gather an initial set of preliminary data on the effects of the probes and contexts, which will guide our subsequent, more rigorous experiments. In total, about 171 ratings from 64 participants were received and analyzed for this preliminary study.

INITIAL RESULTS

Because of the non-independence of data in our preliminary sample, we refrained from inferential tests and focus on descriptive statistics. Also, because the sample sizes within each of the 25 scenario x probe conditions were small, we focus on main effects of probes (across scenarios) and scenarios. This is appropriate because an effective probe should elicit the same threat behavior regardless of the motive of the actor. We are collecting data in a follow-up study that will allow for more thorough inferential analyses.

Table 1 presents the perceived likelihood of insider threat behavior within the different contexts. These results reveal that the social justice and disgruntled worker scenarios generated, on average, the highest likelihood of threat behavior. These groups of workers were particularly more likely than workers in the other scenarios to take detailed notes on products, look for data on servers and computers, and attempt to log into computers. The mean likelihood rating across the six behaviors was 3.44 for the social justice scenario, and 3.37 for the disgruntled worker scenario, indicating a fairly high likelihood of threat behavior under these scenarios. This suggests that social justice and disgruntlement may be important motives for data exfiltration. The next two strongest contexts for data exfiltration were patriotism and new job/financial gain scenarios, with average ratings across the 7 behaviors of 3.19 and 3.00, respectively. The morality/spy scenario yielded the lowest likelihood ratings. The most likely behaviors across all contexts were taking detailed notes (*M* = 3.74), and talking to other researchers about their work (*M* = 3.72). Interestingly, these behaviors would be conducted "off-line" and would operate outside the realm of the technological and performance monitoring tools companies might use to combat insider threats. Logging onto and searching for files on computers and servers were seen as less likely behaviors (*M* = 2.40; *M* = 2.91). The first part of the scenarios depicted what could be considered the 'initial phase' of data exfiltration, wherein the insider may be more concerned with planning than actually searching for data.

Table 1: Mean Likelihood Ratings (and Standard Deviations) of Inside Behavior in Different Contexts (S1: New Job Opportunity; S2: Social Justice (Water); S3: Disgruntled Employee (Quit); S4: Disgruntled Employee (Spy); S5: Patriot)

Actions	Scenarios / Contexts				
	S1	S2	S3	S4	S5
Take detailed notes on his research so that he can bring product information with him	3.7 (1.02)	4.15 (1.17)	4.18 (1.09)	2.72 (1.46)	3.85 (1.33)
Talk to other researchers about their products in informal contexts	3.93 (1.05)	3.85 (1.19)	3.82 (1.24)	3.41 (1.16)	3.59 (1.30)
Look for ways to gather data from servers and other data repositories	2.5 (1.25)	3.33 (1.10)	3.21 (1.34)	2.58 (1.54)	2.97 (1.34)
Attempt to log into different computers to find important data	2.03 (1.00)	2.5 (1.06)	2.67 (1.41)	2.36 (1.37)	2.38 (1.231)
Search for files related to company research on servers or computers	2.8 (1.30)	3.13 (1.14)	2.97 (1.45)	2.59 (1.32)	3.00 (1.35)
Use a camera to take pictures of product designs	2.77 (1.07)	3.53 (1.32)	3.09 (1.49)	2.33 (1.38)	3.03 (1.40)
Download research and product files to a personal USB drive	3.23 (1.25)	3.6 (1.19)	3.59 (1.33)	2.76 (1.48)	3.53 (1.29)

Table 2 presents the mean likelihood ratings for the second part of the scenarios, and tests the effects of the probes. In general, the probes provoked their intended reactions: probes designed to signal an opportunity to exfiltrate data resulted in a higher likelihood of exfiltration behaviors. The probe indicating the presence of a data repository resulted in the highest likelihood ratings (M = 3.32), suggesting that respondents saw this as a good opportunity for the insider to search for, and download data. The probe announcing a forthcoming security software update, and hence the closing of a window of opportunity, resulted in similar average likelihood ratings (M=3.28), suggesting that respondents felt the insider would experience a sense of urgency to respond before the window of opportunity closed. The probes announcing audits resulted in lower ratings for exfiltration behavior, but higher likelihood of removing files unrelate current projects from one's computer. The threat of an audit was perhaps seen as a signal to hide evidence of exfiltration. Interestingly, the probe announcing a brief shutdown of the security software did not promote exfiltration behaviors to the same extent as the probes announcing a new security update and new data repository.

Table 2: Mean likelihood ratings and standard deviations for the impact of probes (email messages) on user actions (P1: Security Update; P2: Weekend Software Upgrade; P3: Training; P4: Security Audit; P5: Backup Server Data)

Actions	Probes				
	P1	P2	P3	P4	P5
Act to secure or obtain confidential data	3.67 (1.24)	3.45 (1.39)	2.92 (1.65)	3.44 (1.50)	3.69 (1.38)
Search for relevant files on servers or computers	3.21 (1.11)	2.88 (1.45)	2.28 (1.34)	3.22 (1.51)	3.59 (1.27)
Attempt to log into different computers to find important data	2.39 (1.14)	2.58 (1.37)	2.33 (1.60)	2.50 (1.18)	2.72 (1.22)
Download research and product files to a personal USB drive	3.61 (1.41)	3.33 (1.51)	2.81 (1.64)	3.25 (1.42)	3.66 (1.43)
Use a camera to take pictures of product designs	3.00 (1.28)	2.97 (1.51)	3.28 (1.65)	2.89 (1.35)	3.09 (1.45)
Remove or delete files for his computer that are not related to his assigned projects	3.79 (1.29)	3.21 (1.58)	3.72 (1.54)	3.20 (1.53)	3.19 (1.45)

CONCLUSIONS

These results suggest that using active probes to stimulate users to act may improve the chances of identifying insider threats. Waiting for naturally occurring events that may trigger a malicious insider to act is both risky and resource intensive, since actual data would be at risk and monitoring will need to be conducted continuously. Our scenario analysis gives us an initial indication that active probes can work for identifying malicious insiders, since our scenarios elicited good response from the subjects. There are differences in response rates based on the specific context, and based on the different probes. This allows us to narrow down the probes for further experimental work, where user responses can be tested in simulated settings. We also plan to use a combination of probes, to determine whether their predictability is increased in combination.

ACKNOWLEDGEMENTS

This paper is prepared with the support of IARPA under Award No. 2016-1603150001. However, any opinions, findings, conclusions or other recommendations expressed herein are those of the author(s) and do not necessarily reflect the views of the IARPA.

REFERENCES

Eldardiry, H., Evgeniy, B., Liu, J., Hanley, J. Price, B. and Brdiczka, O. 2013. *2013 IEEE Security and Privacy Workshops*, San Francisco, CA, 2013, pp. 45-51.

Mazar, N., Amir, O., and Ariely, D.   2008. The dishonesty of honest people: A theory of self-concept maintenance. *Journal of marketing research* (45: 6), pp. 633-644.

McRae, R. R., and Costa, P. T. 1990. *Personality in adulthood*. New York, NY: Guilford Press.

Moore, A., Cappelli, D. Caron, T., Shaw, E., Spooner, D. and Trzeciak, R. 2011. A preliminary model of insider theft of intellectual property. Report by CERT CMU/SEI-2011-TN-Available at: https://resources.sei.cmu.edu/library/asset-view.cfm?assetid=9855

Nu Raytheon Corp. 2009. *Best practice for mitigating and investigating insider threats*. Retrieved on December 6th.

J. R., Buckley, O., Legg, P. A., Goldsmith, M., Creese, S., Wright, G. R., and

Legg, P., Buckley, O., Goldsmith, M., Cresse, S. 2015." *Caught in the act of an insider attack: detection and assessment of insider threat*," presented at the *2015 IEEE International Symposium on Technologies for Homeland Security (HST)*, Waltham, MA, 2015, pp. 1-6.

Whitty, M. (2014, May). Understanding insider threat: A framework for characterizing attacks. *In Security and Privacy Workshops* (SPW), 2014 IEEE (pp. 214-228). IEEE. Chicago.

Schultz, E. E. (2002). A framework for understanding and predicting insider attacks. *Computers and Security*, *21*(6), 526-531.Shaw, E. D., and Stock, H. V. 2011. Behavioral risk indicators of malicious insider theft of intellectual property: Misreading the writing on the wall. *White Paper, Symantec, Mountain View, CA*.

Willison, R., and Warkentin, M. (2009). Motivations for employee computer crime: understanding and addressing workplace disgruntlement through the application of organizational justice. In *Proceedings of the IFIP TC8 International Workshop on Information Systems Security Research. International Federation for Information Processing* (pp. 127-144).