Guest editorial: A brief overview of data leakage and insider threats

Carly L. Huth • David W. Chadwick •
William R. Claycomb • Ilsun You

Published online: 27 February 2013
© Springer Science+Business Media New York 2013

1 Introduction

The challenges of preventing, detecting, and responding to data leakage propagated by authorized users, or *insider threats*, are among the most difficult facing security researchers and professionals today. Prior to the advent of computing, security experts identified potential insider threats by examining suspicious activities in a person's physical behavior. While still relevant in the modern era, we must now also detect suspicious activity in a person's behavior on information systems. But the result is still fundamentally the same: malicious insiders continue to succeed in harming organizations by leaking sensitive information.

Research addressing this problem continues feverishly, but some critical questions remain unanswered. First, can a person's intent be accurately characterized by monitoring and analyzing interactions with computing systems? That is, are the observations made by monitoring and auditing systems robust enough to allow automated characterization of malicious versus non-malicious behavior (or even informed guesses)? Secondly, is the malicious technical behavior of insiders anomalous any more often than the behavior of

non-malicious users? If so, how often are malicious activities clearly anomalous? Finally, is anomalous behavior indicative of potential malicious intent, or do most insiders fall within the boundaries of normal behavior with respect to themselves, their peers, and their organization?

Researchers approach this problem from many different angles. Some propose highly technical solutions, using techniques applied to "Big Data," or various statistical or graph analytic methods. Others attempt to discern the user's intent or disposition via semantic, linguistic, or sentiment analysis of communication such as email or instant messaging. Still others propose combined approaches including analysis of technical events combined with observed behaviors not related to computing systems. Whatever the approach, researchers still struggle to define the problem, much less demonstrate the operational validity of their solutions. In this journal, we document four new approaches seeking to address components of the problem, with the goal of reducing the harm malicious insiders can inflict on an organization.

2 Defining and characterizing insider threats

One of the most important elements in any field of research is the common vernacular researchers use to describe problems and solutions. Unfortunately, insider threat and data leakage research has yet to fully mature in this respect. The literature presents a variety of definitions and characteristics of insiders. These characterizations often focus on different aspects of insider activity, which can be classified as technical, social, or socio-technical approaches to studying insider crime. For example, technical characteristics are the focus of Phygo and Furnell's taxonomy of insider threats, which describes insider activity in terms of network level, system level, and application- and data-level misuses (2004). The social aspects are the focus of Wood's attributes of an insider, which include access, knowledge, privileges, skills, risk, tactics, motivation,

---

C. L. Huth • W. R. Claycomb
CERT Insider Threat Center, Software Engineering Institute,
Carnegie Mellon University, Pittsburgh, USA

C. L. Huth
e-mail: clhuth@cert.org

W. R. Claycomb
e-mail: claycomb@cert.org

I. You (✉)
Korean Bible University, Seoul, Republic of Korea
e-mail: isyou@bible.ac.kr

D. W. Chadwick
University of Kent, Canterbury, UK
e-mail: d.w.chadwick@kent.ac.uk

and process (2000). In contrast, Predd et al.'s four-dimensional approach, describing the organization, individual, system, and environment, presents a socio-technical approach (2008). The CERT Insider Threat Center's current definition of insider threats similarly captures both social and technical elements, focusing on intent, the insider's relationship to the organization and use of information technology (Cappelli et al. 2012):

A malicious insider threat to an organization is a current or former employee, contractor, or other business partner who has or had authorized access to an organization's network, system, or data and intentionally exceeded or misused that access in a manner that negatively affected the confidentiality, integrity, or availability of the organization's information or information systems.

Overall, a variety of approaches are necessary to provide holistic solutions to the problem of insider threats, as illustrated by the variety of methods considered by the papers in this issue. In fact, such diversity was present even in the early days of IT-related insider threat research.

3 Insider threat research overview

In the 1990s, organizations such as the Department of Defense Personnel Security Research Center (PERSEREC) and the RAND Corporation focused academic discourse on the subject of insider threats, through publications and workshops (Anderson et al. 2000; Brackney and Anderson 2004; Anderson 1999; Wood and Wiskoff 2002). Beginning in 2001, the CERT Insider Threat Center also began conducting empirical research into the subject, through the development a database of insider threat cases. During this time period, researchers were also developing models of insider activity, identifying different characteristics of the insiders, the victim organization, and the incident itself. Focusing on behavioral aspects of insider crime, Shaw et al., identified psychological characteristics, such as computer dependency, ethical flexibility and lack of empathy as potential indicators of a risk for destructive and potentially illegal behavior (1998). Focusing on the social-technical approach, Schultz viewed personality traits and technical behaviors, including usage patterns, and meaningful errors, to identify insiders through weighted indicators (2002). Presenting a technical approach, Maybury et al. studied different mechanisms for monitoring technical data in order to detect malicious IT activity (2005).

A recent overview by Hunker and Probst also categorized insider threat studies as technical, social, or socio-technical approaches (2011). In this review, the authors noted that social approaches to insider threats often focus on motivations,

organizational culture, workplace reporting. The authors categorized technical studies as primarily focused in areas such as access controls, monitoring, and policy languages. Finally Hunker and Probst also note work in the socio-technical area, which they note often includes policy, monitoring, prediction, and response work. One example of this approach can be seen in Kandias et al.'s prediction model, which combines psychometric tests and monitoring data from honey tokens and hybrid intrusion detection (2010).

4 Insider threats and data leakage

Data leakage can be characterized as several different types of crimes perpetrated by insiders, including theft of personally identifiable information (to commit fraud, for example), theft of intellectual property, or for an insider to pass sensitive or classified information to an unauthorized third party. It has been noted that data leakage can also refer to inadvertent data loss, however that is beyond the scope of this article (McCormick 2008).

McCormick characterized data leakage and theft into three stages: obtaining access, downloading data, and sharing data (2008). The study outlines common motivation for leaking, often revenge or profit, and notes that leakers sometimes collude with someone outside the organization. McCormick also delineated some potential technical and administrative controls for addressing the threat, including "tightening control on removable media", using data loss prevention tools, and training employees on handling sensitive information.

Other researchers have also addressed the first stage of data leakage, data access. Mathew et al. presented a method of modeling access as a way to begin to understand normal and abnormal access patterns and mitigate insider threats to database information (2010). Aleman-Meza et al. studied the problem of legitimate document access, that is assuring an employee is within their 'need to know' when accessing information (2005). The researchers in this study focused on capturing the context of an employee's need to know and computing semantic associations for documents related to determine the relevance of the document to the employee's need to know.

Surveys and subsequent reports have also been done, providing insight into the amount and impact of insider breaches in a given year. For example, Verizon data breach report includes insiders as a threat agent. The 2012 report noted that internal agents caused four percent of breaches, although they noted that this number might not be representative of all insider data breaches due to low levels of reporting (2012). The report also notes that most internal breaches are deliberate and malicious.

Several studies have been done with a focus on the theft of intellectual property. In 2009, Moore et al. presented two

scenarios for insider intellectual property theft (2009). One model, the Entitled Independent, characterized lone insiders who steal information for a new business opportunity, although in most cases the insider has no specific plans for the information's use. The other model, the Ambitious Leader, characterized insiders who recruit others to steal information, either to develop or benefit a competing organization. More recently, the CERT Insider Threat center has focused on detecting intellectual property theft around the time of employee termination, with Hanley and Montelibano publishing a control and Moore et al. publishing a pattern on the subject (Hanley and Montelibano 2011; Moore et al. 2012). Additionally, Shaw and Stock published a white paper on the psychology of insiders who commit intellectual property theft, noting examples of 'observable workplace risk indicators' and discussing various mitigating strategies, including employee screening and employee reporting programs (2011).

5 Current issue

The papers in this issue address the threat of insider data leakage from a variety of perspectives. Beginning with a behavioral focus, the first paper, "Understanding Insiders: An Analysis of Risk-Taking Behavior" by Fariborz Farahmand and Eugene H. Spafford (2013), explores accepted models of perceptions of risk and the unique characteristics of insider threats. It then introduces metrics to measure the insider's perceptions of risk. In addition, the authors investigate various decision theories, and conclude that prospect theory, developed by Tversky and Kahneman, is the most useful for explaining the risk-taking behavior of insiders.

Moving into more technical solutions, the second paper, "Knowing Who to Watch: Accumulating Evidence of Subtle Attacks" by Howard Chivers et al. (2013), proposes a scalable solution to combining large volumes of evidence from multiple sources gathered over a long period of time. The paper proposes storing long term estimates that entities are attackers, rather than storing the event data itself. The authors identify the essential attributes of the event data and show how to apply Bayesian statistics to update the estimates. They demonstrate the effectiveness of their approach with a simulated slow-attack on a network.

The next article also focuses on a technical approach. Entitled "Two-Stage Database Intrusion Detection by Combining Multiple Evidence and Belief Update" by Suvasini Panigrahi, Shamik Sural and A. K. Majumdar (2013), this article introduces a two-stage database intrusion detection system, which applies anomaly detection for first level inferences followed by misuse detection in the second stage. The system uses inter-transactional as well as intra-transactional techniques for detecting intrusions. The authors analyze the performance of their system using

stochastic models and compare it to two other previously published systems.

The final paper also addresses the protection of databases. "Application of Density-based Outlier Detection to Database Activity Monitoring" by Seung Kim et al. (2013), presents a method for the efficient detection of outliers when monitoring database activity. The authors exploit a kd-tree index and an approximated k-nn search method. The proposed approach was successfully applied to a very large log dataset collected from the Korea Atomic Energy Research Institute (KAERI).

Acknowledgments The guest editors would like to thank the Editors-in-Chief, Professors H. Raghav Rao and R. Ramesh, for providing them with the opportunity to produce this Special Issue. They would also like to extend a special thanks to all the authors and reviewers for the enthusiasm and dedication, without whose help this Special Issue could not have been brought to fruition.

References

- Aleman-Meza, B., Burns, P., Eavenson, M., Palaniswami, D., & Sheth, A. (2005). *An ontological approach to the document access problem of insider threat, proceedings of the IEEE international conference on intelligence and security informatics (ISI) 2005* (pp. 486–491). Georgia: Atlanta.
- Anderson, R.H. (1999). Research and development initiatives focused on preventing, detecting, and responding to insider misuse of critical defense information systems (RAND CF-151-OSD). Technical report, RAND Corporation.
- Anderson, R.H., Bozek, T., Longstaff, T., Meitzler, W., Skroch, M., & VanWyk, K. (2000). Research on mitigating the insider threat to information systems #2 (RAND CF-163-DARPA). Technical report, RAND Corporation.
- Brackney, R.C., & Anderson, R.H. (2004). Understanding the insider threat (RAND CF-196-ARDA). Technical report, RAND Corporation.
- Cappelli, D., Moore, A., & Trzeciak, R. (2012). *The CERT Guide to Insider Threats: How to Prevent, Detect, and Respond to Information Technology Crimes (Theft, Sabotage, Fraud)*. Addison-Wesley Professional.
- Chivers, H., Clark, J. A., Nobles, P., Shaikh, S. A., & Chen, H. (2013). Knowing who to watch: Identifying attackers whose actions are hidden within false alarms and background noise. *Information Systems Frontiers*, 15(1). doi:10.1007/s10796-010-9268-7.
- Farahmand, F., & Spafford, E. H. (2013). Understanding insiders: An analysis of risk-taking behavior. *Informations System Frontiers*, 15(1). doi:10.1007/s10796-010-9265-x.
- Hanley, M., & Montelibano, J., (2011). *Insider Threat Control: Using Centralized Logging to Detect Data Exfiltration Near Insider Termination, Technical Note, CERT*.
- Hunker, J., & Probst, C. (2011). Insiders and insider threats: An overview of definitions and mitigation techniques. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications*, 2(1), 4–27.
- Kandias, M., Mylonas, A., Virvilis, N., Theoharidou, M., & Gritzalis, D. (2010). *An insider threat prediction model, proceedings of the trust, privacy, and security in digital business 7th international conference (TrustBus2010) (LNCS 6264)* (pp. 26–37). Spain: Bilbao.
- Kim, S., Cho, N. W., Lee, Y. J., Kang, S., Kim, T., Hwang, H., et al. (2013). Application of density-based outlier detection to database

- activity monitoring. *Information Systems Frontiers*, 15(1). doi:10.1007/s10796-010-9266-9.
- Matthew, S., Petropoulos, M., Mgo, H., & Upadhyaya, S. (2010). A data-centric approach to insider attack detection in database systems, in Proceedings of Recent Advances in Intrusion Detection: 13th International Symposium, RAID 2010, Ottawa, Ontario, Canada, (LNCS 6307), 382–401.
- Maybury, M., Chase, P., Cheikes, B., Brackney, D., Matzner, S., Hetherington, T., et al. (2005). Analysis and detection of malicious insiders. In Proceedings of the 2005 Intl. Conference on Intelligence Analysis.
- McCormick, M. (2008). Data theft: a prototypical insider threat. In S. Stolfo, S. Bellovin, S. Hershkop, A. Keromytis, S. Sinclair, & S. Smith (Eds.), *Insider attack and cyber security: beyond the hacker* (pp. 52–67). New York: Springer.
- Moore, A., Cappelli, D., Caron, T., Shaw, E., & Trzeciak, R. (2009). *Insider Theft of Intellectual Property for Business Advantage: A Preliminary Model, First International Workshop on Managing Insider Security Threats (MIST 2009)*.
- Moore, A., Hanley, M., & Munide, D. (2012). *A Pattern for Increased Monitoring for Intellectual Property Theft by Departing Insiders, Technical Report, CERT*.
- Panigrahi, S., Sural, S., & Majumdar, A. K. (2013). Two-stage database intrusion detection by combining multiple evidence and belief update. *Information Systems Frontiers*, 15(1). doi:10.1007/s10796-010-9252-2.
- Phyo, A. H., & Furnell, S. M. (2004). *Detection-oriented classification of insider IT misuse, proceedings of the 3rd security conference*. Nevada: Las Vegas.
- Predd, J., Pflieger, S. L., Hunker, J., & Bulford, C. (2008). Insiders behaving badly. *IEEE Security and Privacy*, 6(4), 66–70.
- Schultz, E. E. (2002). A framework for understanding and predicting insider attacks. *Computers & Security*, 21(6), 526–531.
- Shaw, E., & Stock, H. (2011). *Behavioral Risk Indicators of Malicious Insider Theft of Intellectual Property: Misreading the Writing on the Wall, White Paper, Symantec*.
- Shaw, E., Ruby, K. G., & Post, J. M. (1998). The insider threat to information systems: The psychology of the dangerous insider. *Security Awareness Bulletin*, 2, 1998.
- Verizon. (2012). 2012 Data Breach Investigations Report. www.verizonbusiness.com/about/events/2012dbir/. Accessed 6 February 2013.
- Wood, B.J. (2000). An Insider Threat Model for Adversary Simulation, SRI International, Cyber Defense Research Center, System Design Laboratory, Albuquerque, New Mexico.
- Wood, S., & Wiskoff, M.F. (2002). Americans who spied against their country since WorldWar II. Technical Report PERS-TR-92-005, Defense Personnel Security Research and Education Center (PERSEREC).
- Carly L. Huth** is an insider threat researcher in the CERT Program. Huth's current areas of research include the intersections of privacy and technology as well as the effects of the current regulatory environment on insider threat prevention practices. Additionally, Huth participates in assessments of private and public organizations' ability to prevent, detect, and respond to insider threats. Her collaborations with other teams include coauthoring a paper entitled The Impact of Passive DNS Collection on End-user Privacy (SATIN 2012). Huth has experience in both international and academic arenas. Prior to joining the SEI, Huth worked with the Intellectual Property Team in the United Nations Conference on Trade and Development and with the University of Pittsburgh's Office of Technology Management. Huth is a licensed patent attorney and a Certified Information Privacy Professional in Information Technology. She holds a Juris Doctor from the University of Pittsburgh with a Certificate in Intellectual Property and Technology Law. Huth also holds a B.S. from Carnegie Mellon University.
- David W. Chadwick** is Professor of Information Systems Security at the University of Kent. He is the leader of the Information Systems Security Research Group and a member of IEEE and ACM. He has published widely, with over 140 publications, and successfully managed over 25 research projects. He has served as a PC member of over 100 international conferences and been the PC Chair for 5 and co chair for 2. He specializes in distributed policy based systems, Privilege Management Infrastructures, trust management, federated identity management, privacy management and Internet security research in general. He actively participates in standardization activities, is the UK BSI representative to X.509 standards meetings, the chair of the Open Grid Forum (OGF) OGSA Authorisation Working Group, and a member of OASIS and the Kantara Initiative. He is the author of a number of Internet Drafts, RFCs and OGF specifications. His group is the creators of PERMIS, an open source X.509 and SAML supported role based authorisation infrastructure.
- William Claycomb** is the Lead Research Scientist for the CERT Enterprise Threat and Vulnerability Management program at Carnegie Mellon University's Software Engineering Institute. His primary research topic is the insider threat; current work includes discovery of insider threat behavioral patterns and corresponding sociotechnical countermeasures. Dr. Claycomb also works across teams at CERT exploring cloud computing, incident response, systems modeling, and vulnerability analysis. Prior to joining CMU, William was a Member of Technical Staff at Sandia National Laboratories, focusing on enterprise systems security research, including insider threats, malware detection, and data protection. He is currently an adjunct faculty member at CMU's Heinz College, teaching in the School of Information Systems and Management. William received a B.S. in Computer Science from the University of New Mexico, and an M.S. and Ph.D. in Computer Science from the New Mexico Institute of Mining and Technology, where he was the 2010 Patrick Orr Memorial Award recipient.
- Isun You** received his M.S. and Ph.D. degrees in Computer Science from Dankook University, Seoul, Korea in 1997 and 2002, respectively. Since 2005, he has joined Korean Bible University and is currently working as an assistant professor. He served or is currently serving as a general chair or a program chair of international conferences and workshops such as International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS), IEEE International Workshop on International Workshop on Mobility Management in the Networks of the Future World (MobiWorld), IEEE International Conference on Complex, Intelligent and Software Intensive Systems (CISIS), International Workshop on Managing Insider Security Threats (MIST), International Conference on Broadband and Wireless Computing, Communication and Applications (BWCCA) and so forth. He is in the editorial board for International Journal of Ad Hoc and Ubiquitous Computing (IJAHUC), Computing and Informatics (CAI), International Journal of Space-Based and Situated Computing (IJSSC) and Journal of Computer Systems, Networks, and Communications. Also, he served or is currently serving as a guest editor of several journals such as Computing and Informatics (CAI), Wireless Communications and Mobile Computing (WCMC), Intelligent Automation & Soft Computing (AutoSoft), Journal of Intelligent Manufacturing (JIM), Wireless Personal Communications (WPS), Journal of Universal Computer Science (J.UCS), and Ad Hoc & Sensor Wireless Networks (AHSWN) and so forth. His main research interests include internet security, authentication, access control, MIPv6 and ubiquitous computing.