A Pattern for Trust Trap Mitigation

DAVID A. MUNDIE, CERT Program, Software Engineering Institute
ANDREW P. MOORE, CERT® Program, Software Engineering Institute

Insider threat research at the CERT Program has shown that many organizations fall into a vicious cycle of trust and insider threat: organizations do not detect any suspicious insider behavior, so they trust their insiders, do not monitor them, and consequently do not detect suspicious insider behavior. This paper presents a pattern that can break this vicious cycle.

Categories and Subject Descriptors: **H.1.1 [Models and Principles]**: Systems and Information Theory—*Value of information*; **H.1.2 [Models and Principles]**: User/Machine Systems— *Human Information Processing; Human Factors.*

General Terms: Security

Additional Key Words and Phrases: insider threat, security pattern, organizational pattern, enterprise architecture, information security, information assurance, cyber security

ACM Reference Format:

Moore, A.P., and Mundie, D. 2012. A Pattern for Increased Monitoring for Intellectual Property Theft by Departing Insiders. 18th Conference on Pattern Languages of Programs (PLoP). PLoP'11, October 21-23 (October 2011), 7 pages.

1. SUMMARY

As human beings, we naturally vary our scrutiny of others based on the level of emotional trust we have in them. We spend longer studying the appearance of a stranger on our doorstep than that of a good friend. Unfortunately, lessened scrutiny of those we trust can blind us to the precursors of betrayal. This pattern attempts to solve that problem in organizational settings.

2. EXAMPLE

Joe is a mid-twenties high school graduate planning on a career in government service. He takes a job as a clerk in the city traffic court. His pleasant personality and eagerness to please soon ingratiate him with his supervisor, Tracy, and the other staff, and before long he receives administrative access to the traffic fines database and 24-hour access to the courthouse.

Joe is a very social person who likes to hang out with his friends in the evening. One night his friend Tom is furious at the $250 traffic fine he has been given for driving 120 km/h in a 70 km/h zone. Joe casually mentions that perhaps he could take care of the problem for Tom. On his way home, Joe swings by the courthouse, accesses the traffic fine database, and changes the status code on Joe's ticket to "no fine."

Word of Joe's feat spreads, and soon Joe is doing bumper business fixing fines for many of the regulars where Joe and Tom hang out.

3. CONTEXT

Insider threat[1] issues frequently involve two distinct trust domains: intentional and operational[2]. The intentional view focuses on the mental state of individuals, while the operational view involves the actions

Author's address: A.P. Moore, CERT Program, Software Engineering Institute, Carnegie Mellon University, 4500 Fifth Avenue, Pittsburgh, PA 15213; email: apm@cert.org. Michael Hanley, CERT Program, Software Engineering Institute, Carnegie Mellon University, 4500 Fifth Avenue, Pittsburgh, PA 15213; email: mhanley@cert.org. David Mundie, CERT Program, Software Engineering Institute, Carnegie Mellon University, 4500 Fifth Avenue, Pittsburgh, PA 15213; email: dmundie@cert.org.

of individuals. Thus there are two types of trust within an organization. Intentional trust is the aggregate set of emotions and beliefs held by the people in the organization about an insider; it can be measured by asking questions such as “Would you trust Joe with the corporate bank account?” Operational trust consists of the actions of the people in the organization regarding the insider; it can be measured by observing whether, for example, Joe is actually allowed to access the corporate bank account.

There can be a discrepancy between the intentional and the operational domains. Implicit premises that are difficult or impossible to capture in operational terms frequently underpin the intentional domain. For example, when X says that he trusts Joe to “edit the database,” he really means “edit the database for the benefit of the company,” which is very difficult to express in an operational system.

More specifically, the current pattern is applicable to IT-intensive organizations with a traditional management structure, a mature IT staff, a culture of process improvement, and resources sufficient for addressing insider threat issues.

4. CONTEXT

Too often organizations fall into the “trust trap” [Band 2006], which says that as intentional trust grows, operational trust should grow likewise. On the surface, this approach has an appealing, common-sense feel to it: surely the more good feelings management has about individuals, the less those individuals need to be monitored and the fewer controls need to be applied to them. It seems to improve morale by showing insiders that they are respected and valued; maximize organizational resources by not deploying unnecessary monitoring; and let management feel that they have created a mature, well-run organization that does not rely on sanctions to function efficiently.

However, despite its intuitive appeal, this approach engenders an important weakness: a vicious cycle involving monitoring and intentional trust. The less an organization monitors an employee, the fewer opportunities there will be to detect behavior that would diminish the intentional trust placed in the employee, which leads to even less monitoring. Observing fewer bad behaviors leads the organization to trust the employee more and monitor him or her even less. Combined with a frequently concomitant loosening of the access controls applied to the employee, and with the fact that decreased monitoring can embolden employees contemplating malicious activities, the trust trap exposes the organization to the possibility of a serious insider attack. The current pattern helps organizations balance operational and intentional trust to avoid the trust trap.

Figure 1 is a causal loop diagram[3] [Anderson 1997] that illustrates the trust trap. The loop on the left is a reinforcing loop: the initial intentional trust of the insider causes the monitoring of the insider to decrease, which means that suspicious acts are more likely to go undiscovered, which allows the intentional trust of the insider to go up, all in a self-reinforcing pattern. The loop on the right is a stabilizing loop: the monitoring of the insider leads to the discovery of suspicious acts by the insider, which investigation may reveal to be actual malicious or criminal acts, which decreases the number of insider malicious acts through either the termination or the punishment of the insider.

[1] The CERT program defines the term “insiders” as “Current or former employees or contractors who intentionally exceeded or misused an authorized level of access to networks, systems, or data in a manner that targeted a specific individual or affected the security of the organization’s data, systems, and/or daily business operations.”

[2] Interestingly, these two domains correspond to the two major psychological schools of the twentieth century: cognitivism, which focuses on mental processes, and behaviorism, which focuses on the actions people take.

[3] In causal loop diagrams, arrows represent the pair-wise influence of the variable at the source of the arrow on the variable at the target end of the arrow. An arrow labeled S indicates that the values of the variables move in the same direction, whereas an arrow labeled O indicates that they move in the opposite direction.

A powerful tenet of system dynamics causal loop diagrams is that a problematic behavior’s underlying feedback structure captures the behavior’s dynamic complexity. Causal loop diagrams identify two types of feedback loops: balancing and reinforcing. Reinforcing loops describe system aspects that tend to drive variable values consistently upward or downward and are often typified by escalating problematic behaviors. Balancing loops tend to drive variables to some goal state and are often typified by aspects that control problematic behaviors.

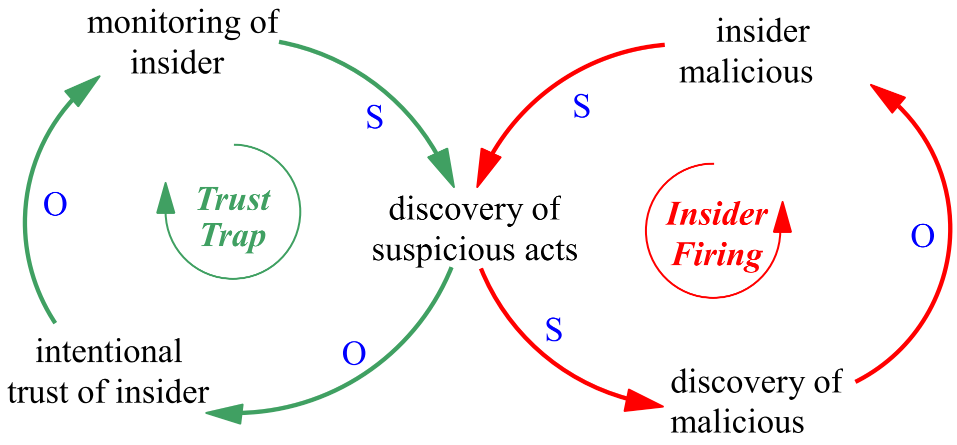

Fig. 1. A causal loop diagram of the trust trap

A solution must resolve the following forces:

- the need to monitor employees against the desire to show that they are respected and valued
- the beneficial effects of early detection of malicious insider activity against the monetary and opportunity costs of a monitoring program
- the need for fairness and disgruntlement mitigation against the complexities of designing a monitoring program

5. SOLUTION

The trust trap is so widespread in part because adjusting operational trust to match intentional trust does, in fact, work most of the time and can be an effective heuristic. There is even a biological basis for its effectiveness. Recent research has demonstrated a neurological basis for the interplay of trust and trustworthiness. Individuals who perceive that they are trusted show a rise in their levels of oxytocin, the so-called social binding hormone. This rise in oxytocin is in turn associated with more trustworthy behavior in those individuals [Zak 2004, Zak 2008].

It is rare that an employee you feel is trustworthy will betray you. It is essentially impossible to completely abandon the heuristic and apply the same degree of monitoring to trusted insiders as to outsiders. However, there are a number of activities that can help balance the forces.

Maintain morale. The potential adverse effects of employee monitoring on morale can be mitigated by two measures. The first is effective communications with employees about monitoring and why it is done. It is unlikely that employees in North America will come to be flattered by monitoring as is the case in some Asian countries, but a well-planned awareness campaign can go a long way.

The second measure is to make sure that the monitoring is scrupulously fair and applied impartially. Unfairness has been shown to be an important component of disgruntlement, which in turn has been shown to be a major component of insider threats.

Optimize resources. To minimize the cost of an employee monitoring program, three techniques have proven useful. The first is to use spot checks rather than continuous monitoring. The second is to avoid distributing the details of the monitoring activities so that employees assume that there is more monitoring than there actually is. The third is to increase monitoring at times of greatest risk, such as during the last month of employment (see “A Pattern for Increased Monitoring for Intellectual Property Theft by Departing Insiders”).

Educate management. We all like to be loved, and scrutinizing employees to detect signs of betrayal is something that many managers are loath to do. Here too an awareness program that makes clear that charisma and bonhomie are not enough, and that good managers must be prepared to sacrifice some of their congenial aura to protect the assets of the organization.

Monitor accurately. Given the current state of technology, it is usually impossible to express intentional policies in operational terms. The intentional policy may be something like “Alert me whenever Mr. Smith sends suspicious e-mail to Uruguay,” but implementing that would require complex programming, simplifying assumptions, and questionable approximations. The pattern “Early Detection of Insider Activity” is the framework for improving detection, but two specific best practices are recommended here. The first is to use a formal language such as Description Logic (DL) [Breaux 2008] to express policies. Even if that language is not processed automatically, it will at least give a concise, unambiguous version of the policy. The second is to start from a clean, unambiguous model of insiders within the organization, using something like Gates and Bishop’s lattice model [Bishop 2008] of groups, assets, and controls, as this gives a clean framework that is tailored for insider threats.

Secure results. The monitoring infrastructure, and especially the data collected, are highly sensitive and need to be protected by suitable authentication and authorization mechanisms

6. STRUCTURE

Figure 2 shows the structure of the solution. The trust trap mitigation steps have increased the monitoring of the insider so that the organization can detect his or her suspicious acts, resulting in decreased intentional trust of the insider.

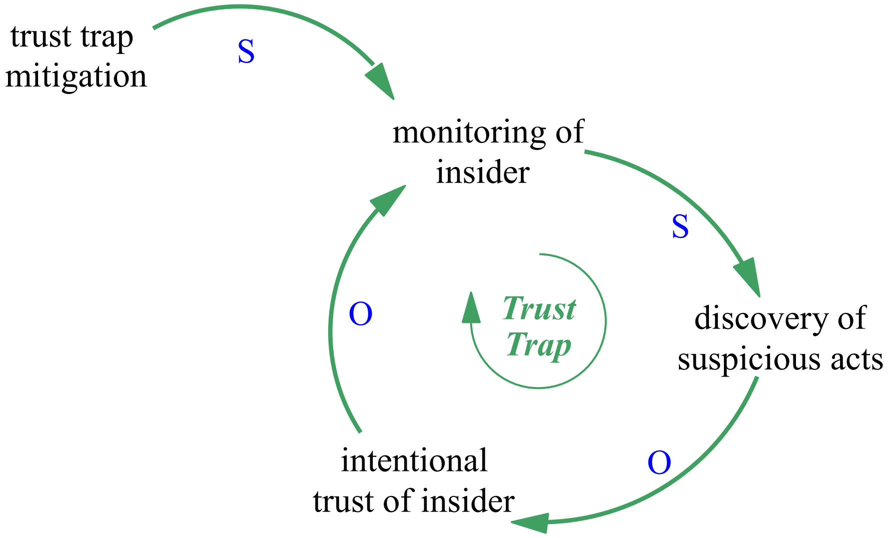

Fig. 2. The stabilized trust trap

7. DYNAMICS

The dynamics of the solution are conceptually simple, although they can seem cumbersome to management. Following established policies and procedures, the organization’s entire staff—the IT staff especially—monitors what goes on with an eye to detecting events that could be malicious activity or its precursors. If staff detect suspicious events, they report them through the organization’s documented reporting process. This activates the right-hand loop in Figure 1, lowering the organization’s trust in the individual involved. In extreme cases, that diminution in trust can result in the termination of the insider.

8. IMPLEMENTATION

The following checklist is useful when using this pattern.

a. Understand your legal obligations and responsibilities regarding privacy. If at all possible, obtain employee consent for monitoring (see [Huth 2011]).

b. Review policies to make sure you give employees a clear general overview of employee monitoring within your organization.

c. Determine your budget for employee monitoring and what level of monitoring fits within that budget.

d. Design a monitoring program that is flexible enough to respond to changes in your trust level and to allow easy spot checks and increased granularity.

e. Map your high-level monitoring design onto low-level rule sets to implement it.

f. Establish a threat model and use it to define what situations warrant increasing the level of monitoring.

g. Have your monitoring plan reviewed by an ethics office, by human resources, or by a legal representative to ensure its fairness.

h. Plan an awareness campaign for managers that explains their role in avoiding the trust trap.

9. EXAMPLE RESOLVED

Despite the emotional trust Joe's manager Tracy has in him, she requests daily reports from her physical security staff of any unusual after-hours comings and goings. So the day after Joe fixes Tom's ticket, Tracy calls Joe into her office and asks him what he had been doing at the office so late the day before. Joe flushes bright red and stammers that he had realized he had made an error in the database and simply could not rest until it was fixed.

Tracy had reviewed the database logs, but the log was not fine-grained enough to show exactly what Joe had done, only that he had accessed the database. Tracy's suspicions are aroused, and she chides Joe about his after-hours action, gently making the point that she has her eye on him. She has IT increase the granularity level on the database log and increase the monitoring level on Joe's activities for a two-month period.

Joe, embarrassed at having been caught and now fearful of Tracy's monitoring, decides it is not worth jeopardizing his career just to appear to be a big shot in front of his friends. He continues to work in the court office another five years with no further incidents.

10. KNOWN USES

The authors are unaware of any implementation of the pattern in a production environment. However, the individual components of the solution—employee monitoring, spot checking, hiding the details of monitoring, risk-based monitoring, management education, awareness programs, formal policy specifications, and securing the results of monitoring—are widely deployed. Further, most of them are embodied in cyber-security standards such as the Federal Information Security Management Act of 2002 (FISMA), ISO/IEC 27002:2005, and the CERT® Resilience Management Model. For example, NIST Special Publication 800-53 specifies security controls for FISMA and incorporates best practices for position categorization and personnel screening (PS-2 Position Categorization and PS-3 Personnel Screening), maintenance of explicit rules of behavior and consequences for violation (PL-4 Rules of Behavior and PS-8 Personnel Sanctions), audit log monitoring (AU-6 Audit Review, Analysis, and Reporting), and security awareness and training (AT-2 Security Awareness and AT-3 Security Training). The requirement that controls must be based on broad community consensus prior to their inclusion in NIST 800-53 and the requirement that federal agencies must by law implement these controls indicate that the basis for the Trust Trap Mitigation pattern is tried and true.

® CERT Resilience Management Model is registered in the U.S. Patent and Trademark Office by Carnegie Mellon University.

Random spot checking is required by federal policy. Hiding the details of monitoring is a widely used practice in the intelligence community. Increased monitoring during periods of increased risk is an obvious strategy, recently embodied by Hanley [2010]. Cranor [2002] exemplifies formal specification of privacy policies, and Breaux [2008] exemplifies privacy regulation.

11. CONSEQUENCES

The net result of this pattern is the moderation of the reinforcing trust trap loop via inputs that increase monitoring. The primary mechanism is management awareness of the problem and willingness to perform the monitoring, but other mechanisms, such as fairness and minimal resource expenditure, play an important role in diminishing unintended consequences of such monitoring.

This pattern results in the following benefits:

Constant verification of emotional and behavioral trust. The balancing of emotional and behavioral trust is not a new problem. “Trust is good, but control is better,” said Vladimir Lenin, while Ronald Reagan was pithier: “Trust, but verify.” The current pattern avoids the extremes of paranoia and gullibility.

Continued demonstrations of respect for employees. By including employees in the monitoring design process, and by educating managers in disgruntlement mitigation strategies, this pattern enables the organization to have its cake and eat it too: the organization can obtain the data it needs while maintaining the good will of its employees.

Optimal use of monitoring resources. Monitoring can be a very expensive proposition: the quantity of data is enormous; the analysis of that data is an AI-complete problem, meaning that it cannot be truly solved without strong artificial intelligence; and high false positive rates are endemic. The strategies in the current pattern alleviate this problem.

Employee acceptance of the need for monitoring. The goal of this pattern is not just to avoid disgruntlement, but also to enlist employees in the enterprise-wide insider threat mitigation program.

Protection of organizational assets. By increasing the level of detection, this pattern contributes to decreasing the organization’s exposure to theft of intellectual property, fraud, and sabotage.

This pattern suffers from the following drawbacks:

Complex monitoring program. Designing a monitoring program that balances the forces is not easy, and evaluating its effectiveness is, as with any security measure, problematic.

Continuing gap between monitoring policy and its implementation. Given the current state of technology, it is not possible to specify exactly what events should trigger alerts, let alone to implement those specifications exactly.

12. SEE ALSO

The parent pattern for the Trust Trap Mitigation pattern is the “Early Detection of Insider Activity” pattern, which stresses the need to monitor insider attack precursors to detect actual attacks as early as possible. A child pattern of the Trust Trap Mitigation pattern is “A Pattern for Increased Monitoring for Intellectual Property Theft by Departing Insiders,” which optimizes employee monitoring by reinforcing it during periods of increased risk, such as just prior to insider termination. A sibling pattern is the “Rapid Escalation” pattern for making sure that alerts are handled in a timely fashion. All of these patterns will appear in Moore [2011].

13. EPILOGUE

This pattern is part of the insider threat pattern system being developed at the CERT Program. The goal of this research is to enable coherent reasoning about how to design and, to a lesser extent, implement enterprise systems to protect against insider threat. Instead of being faced with vague security requirements and inadequate security technologies, system designers will be armed with a coherent set of architectural patterns that will enable them to develop and implement effective strategies against the insider threat in a timelier manner and with greater confidence.

14. ACKNOWLEDGEMENTS

The authors extend their heartfelt appreciation to Dr. Eduardo Fernandez, professor in the Department of Computer Science and Engineering at Florida Atlantic University. His review and guidance as part of the PLoP 2011 conference shepherding process led to immense improvement in the understanding and articulation of this pattern. In addition, we thank Paul Ruggiero of the Software Engineering Institute for his careful and detailed technical editing of this pattern, which dramatically improved its readability and flow. Finally, thanks to Carly Huth for providing helpful advice regarding the legal aspects of employee monitoring.

REFERENCES

Anderson, V., Johnson, L.: Systems Thinking Basics: From Concepts to Causal Loops. Pegasus Communications (1997)

Band, S.R., et al.: Comparing Insider IT Sabotage and Espionage: A Model-Based Analysis. SEI Technical Report CMU/SEI-2006-TR-026 (December 2006)

Bishop, M., Gates, C.: Defining the Insider Threat. In: Proceedings of the Cyber Security and Information Intelligence Research Workshop, article 15 (May 2008)

Breaux, T.D., Antón, A., Doyle, J.: Semantic Parameterization: A Process for Modeling Domain Descriptions. ACM Trans. Softw. Eng. Methodol. 18, 2, 1–27 (2008)

Cranor, L.F.: Web Privacy with P3P. O'Reilly & Associates (2002)

Hanley, M.: Candidate Technical Controls and Indicators of Insider Attack from Socio-Technical Models and Data. In: Proceedings of the 2010 NSA CAE Workshop on Insider Threat (2010)

Huth, C.: Monitoring and Privacy: A Legal Perspective. Unpublished.

Moore, A., et al.: Architecting the Enterprise to Protect Against Insider Threat. Unpublished.

Zak, P.J., et al.: The Neurobiology of Trust. Ann. N.Y. Acad. Sci. 1032, 224–227 (2004)

Zak, P.J.: The Neurobiology of Trust. Scientific American (May 2008)