The Validity of Penetration Tests

Penetration (pen) tests are critical to operating and maintaining an effective information security program. They are used for a variety of purposes, including assessing system readiness, identifying gaps, assigning resources and evaluating vendor viability. These tests are important, but how do reviewers establish credible results on which to base decisions? Should results be taken at face value? What external effects influence findings? This article contends that external factors, such as compliance and market pressure, can affect, and do detract from, the validity of penetration test results.

Because of the value and prevalence of compliance frameworks, such as the Payment Card Industry Data Security Standard (PCI DSS) and the US Federal Risk Authorization Management Program (FedRAMP), enterprises may be under the impression that achieving compliance is enough. This belief may lead to a false sense of security.[1] Although some examples in this article are in the context of PCI and FedRAMP frameworks, the guidance is applicable to all penetration tests. Suggestions from this article apply to any penetration test, because the goal of these tests is to promote security awareness and conduct critical source evaluation.

The Value in Pen Testing

A security professional's understanding of how an information system operates versus what the system really does is likely different. Rob Joyce, former chief of the US National Security Agency Tailored Access Operations (NSA TAO), best summarized this premise:

> *You know the technologies you intended to use in that network. We know the technologies that are actually in use in that network. Subtle difference. You'd be surprised about the things that are running on a network vs. the things that you think are supposed to be there.*[2]

These unintended discrepancies within systems serve as the exploitable vectors that can be leveraged by malicious actors to gain access to systems and compromise data. Physical security and documented policies and procedures provide the benefit of tangible constructs, lending to simple verification through observation. Because these tangible constructs are directly observable, they are straightforward to identify and remediate. An unlocked or defective door is plainly observable and verifiable. Similarly, an enterprise understands physical inventory and boundary, because these also are casually observable. The equivalent of these physical checks becomes increasingly complex when it is a logical abstraction at the technical level of an information system.

Continuing the analogy of the door, consider Active Directory (AD). Active Directory is a special-purpose database that is used to manage logical objects and attributes,[3] and, similar to a physical door, enforce access controls. At what point is the AD installation verified as being secure? Enterprises do not intentionally configure insecure AD instances or lose them; however, these errors occur. Poorly installed physical security is easily observable. Conversely, poorly installed logical abstractions are unobservable. Simply because an abstraction is installed does not mean that it operates as intended in the context of the larger system. The ease with which flaws are introduced to an information system stresses the significance of penetration tests.

Penetration testing offers a solution to the problem of logical verification of a system's security via purposeful controlled attack. Penetration testing

Brent Michel, CISA
Is a veteran of the US Air Force and a professional penetration tester for Coalfire Labs, the technical business unit of Coalfire Systems Inc. Michel conducts network, application and physical penetration testing against Fortune 500 information systems. Formerly, he advised secure program development for US federal cloud systems under directives for controlled unclassified information.

consists of the reconnaissance and exploitation of vulnerabilities in hardware and software caused by misconfiguration and user behavior. Penetration testing is a form of alternative analysis. "The objective of alternative analysis is to hedge against natural human and organizational constraints using liberating structures or structured analytic techniques, or by employing a wholly different team not already immersed in an issue to challenge assumptions or present alternative hypotheses or outcomes."[4]

Threats to Validity

One of the challenges that any penetration test faces is demonstrating the validity of the testing methodology, execution and operator to ensure that the results can be trusted. It is important to assess the validity of any penetration test that is used to support change management, whether evaluating a vendor, payment processor, appliance, cloud service or simply trying to understand the real-world posture. Good research is valid research, and valid research means conclusions drawn are legitimate.[5] The practice of restricting scope to demonstrate favorable results threatens the external validity of penetration testing. These scope limitations are put in place for many reasons, such as:

> **"One of the challenges that any penetration test faces is demonstrating the validity of the testing methodology, execution and operator to ensure that the results can be trusted."**

- Compliance is a business enabler. Compliant information technology solutions are allowed access to markets otherwise unobtainable.
- A conflict of interest exists for enterprises that wish to access these markets, but do not value security or compliance.
- The enterprise being audited is responsible for selecting and compensating the auditor organization. This can remove independence of the auditor.
- Information security lacks the maturity of other critical industries.[6]

Compliance as a Business Enabler

Compliance allows businesses the opportunity to interact with markets that they otherwise could not. A prime example is the US Federal Cloud Computing Strategy, which led to the creation of FedRAMP. The Federal Cloud Computing Strategy set the initiative to migrate aging US federal systems to cloud platforms with the benefit of cost savings. At the time that the Federal Cloud Computing Strategy was authored, the federal government maintained a US $80 billion IT budget. Subsequently, FedRAMP opened a US $80 billion market to the private sector that had previously been unobtainable. Today, that budget is passing US $90 billion.[8]

"Information security has become such an integral aspect of being able to build brand and advance progress."[9] IT providers have taken notice. Look no further than examples of FedRAMP in-process solutions,[10] such as Amazon Web Services, Microsoft and VMware compliance pages.

Independence and Conflict of Interest

The need for independent auditing derives from the fact that an enterprise cannot grade its own work. To simplify the dilemma to a grade-school metaphor, picture the capability of students to grade their own school work. The students cheat and give themselves 100 percent on each assignment. This works well until the final exam, for which the instructor explains that students may no longer grade their own work. "This would expose all the things that you should have learned or maybe thought you understood, but never really did. Grading your own homework might feel good in the short term, but it completely clouds one's awareness and can eventually lead to a failing grade."[11] Grading one's own school work is like auditing oneself; it looks good, until challenged.

Enjoying this article?

- **Learn more about, discuss and collaborate on cybersecurity in the Knowledge Center.** *www.isaca.org/cybersecurity-topic*

When an enterprise selects and compensates its auditor, the same result (advertently or inadvertently) as grading one's own school work can be accomplished. Ideally, the organization responsible for upholding the standard would be the organization that performs the audits. Although compliance is valued, the accompanying security measures required for compliance may pose an obstacle whose benefits are difficult to justify.[12] In instances where an enterprise desires certification, but not the accompanying security, the audit may be undermined.[13]

Undermining Pen Tests Results

Penetration tests are "inevitably influenced by a company's motivation for subjecting itself to one. When compelled by regulators or insurers, it is usually doomed to be a perfunctory, check-the-box exercise, scoped as narrowly as possible."[14] This narrowed scope occurs for a few reasons:

- Scope reduction reduces the attack surface and minimizes the likelihood of identifying discrepancies.
- The auditor/organization does not understand the system inventory or system operation.[15]
- The auditor/organization realizes scoping inaccuracies after the penetration test is conducted.

The term "auditor" in the previous bulleted list is intended to reference the individual guiding an audit, who is not always within the same enterprise and who is not the penetration tester. For scoping purposes, it is better to avoid having the auditor establish hard scope boundaries of the penetration test. Doing so could result in precluding viable real-world attack vectors. In the PCI, an enterprise hires a qualified security assessor (QSA). The enterprise, along with the QSA, determines the scope of the assessment. The scope is then contracted for penetration testing. FedRAMP shares similarities.

Penetration tests are neutral third-party observations within a predetermined scope. Professional experience in this area has observed trends where scope limitations intentionally or unintentionally omit critical assets governing access controls or even restrict basic methods. Another observation is that engagement timelines limit full network compromise. In a drastic instance, two Internet Protocol (IP) addresses were provided as an indicative measure for the security of an entire cloud infrastructure. Stories are frequent within the penetration testing community of confused enterprises that are failing to see how their accredited systems (having undergone compliance pen tests) can be compromised when scope is expanded during another penetration test. Narrowly scoped engagements are not realistic adversarial assessments of the ability of any information system to withstand or repel an attack. Regardless of severe limitations, providers in these scenarios are still gaining accreditation. This is largely due to information security being very much in its infancy compared to its older industrial peers, such as aviation.

> **“Narrowly scoped engagements are not realistic adversarial assessments of the ability of any information system to withstand or repel an attack.”**

Information Security as a Developing Field

Information security is a maturing field, especially when compared to aviation. Much like aviation, certification programs must test whether equipment "can in fact be operated by people with the level of skill and experience presumed in the specification."[16] In enterprises like the US Air Force, it is normal to establish multiple levels of quality assurance to prevent and identify the cause of defects.[17]

Within the precision measurement field of aviation, measurement traceability is established to international standards. Metrologists audit individual pieces of equipment against predetermined specifications. The metrologist who is conducting the audit is certified annually by a quality assessor who observes the calibration process in its entirety. Apart from the certification process, metrologist output is sampled and independently verified by the quality assessors at a rate of 5 percent. Discrepancies in output between the metrologist and the quality assessor elevate to a root-cause analysis, during which the metrologist must re-perform the audit to demonstrate technique. Biannually, the US Air Force observes metrologists and quality assessors to ensure that both meet program standards. Deficient laboratories are stripped of certification, and past work is reaudited after extensive training.[18] After 100 years of aviation, it has some of the best practices established to help prevent loss of life. No information security compliance framework exists today that holds equivalent standards.

Implications for Security Managers

Logical security is tough to verify without actionable penetration tests; this is where penetration tests derive their value. Like any research, the dilemma is discerning valid from invalid pen test results. Like many issues within security, awareness of the problem is a significant first step to remedying the issue. Understanding the shortfall of compliance practices provides practitioners with incentive to critically evaluate information systems prior to accepting attestations at face value. This is important for any practitioner who is concerned with the security of their information, particularly as they partner with external vendors, systems or technology, or evaluate their own posture.

The following are questions that practitioners can ask the next time a pen test comes across the desk:

1. Was the pen test conducted in an effort to sell something?[19]
2. What was the scope of the engagement?
3. What was scoped out from the engagement? Why?
4. How many hours were directed toward the actual test? Is there a possibility a system is deemed secure because time precluded a full compromise?
5. How skilled are the individuals who are conducting the test?
6. Did any of the pen test findings not require remediation? Why?
7. Are there types of attacks that were not part of testing (e.g., denial of service)?
8. Was testing done in a production or staging environment?
9. Did the environment house data, and were data being actively transported during the time of the test?

> **“Practitioners should never accept penetration test results at face value. Like all reports, the results can be manipulated to meet a predetermined outcome.”**

Practitioners should never accept penetration test results at face value. Like all reports, the results can be manipulated to meet a predetermined outcome. All sources that are attesting a secure stance should be critically evaluated.

Author's Note

The opinions expressed in this article are personal. The content is not monitored or approved by Coalfire Systems Inc. and does not represent the views and opinions of Coalfire Systems Inc.

Endnotes

1 Thurman, M.; "Compliance Does Not Equal Security," *ComputerWorld*, 12 January 2016, *www.computerworld.com/article/3021787/security/compliance-does-not-equal-security.html*

2 Zetter, K.; "NSA Hacker Chief Explains How to Keep Him Out of Your System," *Wired*, 28 January 2016, *https://www.wired.com/2016/01/nsa-hacker-chief-explains-how-to-keep-him-out-of-your-system/*

3 Microsoft Press, *Introducing Windows 2000 Server,* 2000, *https://msdn.microsoft.com/en-us/library/bb742424.aspx*

4 Zenko, M.; *Red Team*: *How to Succeed by Thinking Like the Enemy,* Basic Books, USA, 2015

5 Stangor, C.; *Research Methods for the Behavior Sciences,* Wadsworth, USA, 2011

6 Anderson, R. J.; "Why Cryptosystems Fail," *Communications of the ACM,* 1994

7 Kundra, V.; "Federal Cloud Computing Strategy," The White House, USA, 8 February 2011, *www.whitehouse.gov/sites/default/files/omb/assets/egov_docs/federal-cloud-computing-strategy.pdf*

8 Moore, J.; "5 Charts That Show Where the Federal IT Budget is Headed Next Year," *Nextgov*, 8 March 2016, *www.nextgov.com/cio-briefing/2016/03/5-charts-show-where-federal-it-budget-headed-next-year/126520/*

9 Siwicki, B.; "Chief Information Security Officer Salaries Skyrocket as High as $420,000, Survey Finds," *Healthcare IT News*, 24 May 2016, *www.healthcareitnews.com/news/chief-information-security-officer-salaries-skyrocket-high-420000-survey-finds*

10 FedRAMP.gov, "FedRAMP At A Glance," 21 August 2016, *https://marketplace.fedramp.gov/#/products?sort=productName*

11 *Op cit,* Zenko

12 Palmer, D.; "Cyberattacks Are Growing, but Still Nobody Knows What They Really Cost," *ZDNet*, 11 August 2016, *www.zdnet.com/article/cyberattacks-are-growing-but-still-nobody-knows-what-they-really-cost/*

13 Moore, D. A.; P. E. Tetlock; L. Tanlu; M. H. Bazerman; "Conflicts of Interest and the Case of Auditor Independence: Moral Seduction and Strategic Issue Thinking," *Academy of Management Review*, vol. 31, no.1, January 2006, *http://citeseerx.ist.psu.edu/viewdoc/*download?doi=10.1.1.576.6577&rep=rep1&type=pdf

14 *Op cit,* Zenko

15 *Op cit,* Anderson

16 *Ibid*.

17 United States Air Force, Air Force Instruction 21-101, 21 May 2015, *http://static.e-publishing.af.mil/production/1/af_a4/publication/afi21-101/afi21-101.pdf*

18 United States Air Force, Air Force Metrology and Calibration Program, 2011

19 Ede, L.; *The Academic Writer: A Brief Guide*, Bedford Books, USA, 2014