Evidence-Based, Good Enough, and Open

Karen Scarfone
Computer Security Division
National Institute of Standards and Technology (NIST)[1]

A profoundly important question to organizations is "how secure are our systems?" However, organizations cannot answer this question with any degree of certainty. We are developing a new approach to answering this question. This approach is distinguished from previous efforts in three ways:

1. **Evidence-based.** Security decision-making should not be based on conventional wisdom, but instead on enhancing threat models and risk assessment methodologies so that they leverage the results of analyzing historical and current operational and technical metrics related to vulnerabilities, attacks, and security controls.
2. **"Good enough" answers.** Ultimately, it should be possible to precisely measure the security of a system. However, while long-term research is underway to determine how this can be accomplished most effectively, there is still an immediate need to have measurements that could be used to assist organizations in making sound security decisions. Our approach focuses on developing the best measurements that we can in the short term, with the understanding that more precise measurements will replace them in the long term.
3. **Open specifications and standards.** The Security Content Automation Protocol (SCAP) work combines several open specifications for expressing and collecting security data. These specifications are intended to help automate vulnerability and patch management, verify compliance with security policies and baselines, and score the relative severity of vulnerabilities.[2] All of these use cases involve security data that may be used for metrics, so we encourage use of these interoperable specifications for expressing, collecting, and analyzing security measures and metrics, and we plan on developing additional specifications as needed. Specifications are being developed by the MITRE Corporation, the Forum of Incident Response and Security Teams (FIRST), and NIST.

Our approach is a new framework to support evidence-based quantitative risk assessment. It is intended to enhance, not replace, existing threat modeling and risk assessment methodologies. The framework has several applications. In addition to measuring the overall security of a system, it can also be used to compare the relative security of different systems and to measure the likely security impact of a change to a system or its environment. Information collected through the framework could also be used as evidence for compliance evaluations (e.g., Health Insurance Portability and Accountability Act [HIPAA], Sarbanes-Oxley [SOX], Federal Information Security Management Act [FISMA]).

---

[1] Official contribution of the National Institute of Standards and Technology; not subject to copyright in the United States.

[2] http://nvd.nist.gov/scap.cfm

We are currently in the early stages of what we expect to be a five-year project. We recently completed the initial planning of the framework and documented in a paper presented at the IEEE 1st International Conference on Information Technology in May 2008.[3] We have also been studying how the Common Vulnerability Scoring System (CVSS) could be applied to security configuration issues, and we published a draft specification for the new Common Configuration Scoring System (CCSS) for public review and comment in May 2008.[4] We have also been analyzing CVSS scores from the National Vulnerability Database (NVD) to better understand the typical characteristics of CVSS scores and to find possible improvements for CVSS.[5]

The following is a list of some of the questions that we will be attempting to answer through our research. We are particularly interested in hearing attendees' opinions on these questions and in identifying any other work in this area.

- Which characteristics of vulnerabilities and attacks are most relevant for predicting the likelihood of future compromises?
- Which types of scoring scales would be most useful? Examples are numeric scales (e.g., 0-10, 0-100), mean time to exploitation, and rankings. Multiple scales may be needed to accomplish different purposes.
- How should low-level scores for individual vulnerabilities and attacks be rolled up into host security scores? Can host security scores be rolled up into network or enterprise security scores? If so, how?
- Which classes of vulnerabilities should be considered? Software flaws, security configuration issues, software feature misuse, others?

---

[3] Scarfone, K. and Grance, T., "A Framework for Measuring the Vulnerability of Hosts", *Proceedings of the 1st International IEEE Conference on Information Technology*, Gdansk, Poland, May 2008.

[4] Scarfone, K. and Mell, P., NIST Interagency Report 7502: *The Common Configuration Scoring System (CCSS) (DRAFT)*, May 2008. http://csrc.nist.gov/publications/PubsNISTIRs.html

[5] Mell, P. and Scarfone, K., "Improving the Common Vulnerability Scoring System", *IET Information Security*, September 2007.