Measure Like You Meant It

DAN GEER AND RICHARD BEJTLICH

Dan Geer is the CISO for In-Q-Tel and a security researcher with a quantitative bent. He has a long history with the USENIX Association, including officer positions, program committees, etc.
dan@geer.org

Richard Bejtlich is chief security strategist at FireEye and a nonresident senior fellow at the Brookings Institution.
contact@taosecurity.com

"How did it get so late so soon?"

—Dr. Seuss

Pie charts: managers of security programs create them; their bosses consume them. What slice of corporate computers is patched? What percentage of systems runs antivirus software? What versions of various Microsoft Windows operating systems are present in the environment? Pie charts are the answer to questions like those.

The only real purpose of security metrics is decision support; therefore, we question the utility of pie charts. Statistics on patching, antivirus, and OS distribution are needful, but remember—they are "input metrics." They are a means to an end; they are not the end itself.

Managers devise and operate digital security programs in many forms, sometimes summarizing them in the form of the confidentiality-integrity-availability triad. One goal of a digital security program might well be to protect the organization's data from theft by an intruder. Assuming for the sake of argument that such protection is the organization's primary goal, might it not make sense to measure progress toward that goal?

Organizations may well care about mitigating data theft, but seldom do they invest in measuring their progress toward that goal, per se. They spend more time (creating pie charts) on input metrics like patching, antivirus, and OS, but they don't track "output metrics." What they need to ask is, "Are we compromised, and, if so, how bad was the intrusion?" You can spend all the time in the world measuring what goes into the oven, but that won't tell you if you burned the cake—or if a thief stole it.

We argue that the two output metrics for understanding the risk of data theft are (1) counting and classifying digital security incidents, and (2) measuring the time elapsed from the moment of detection to the moment of risk reduction.

1. **Counting and classifying digital security incidents:** An organization should devise a tracking mechanism appropriate for its environment and culture. Security professionals are sure to debate the nature of various intrusions, but don't allow that debate to drag on; you need a taxonomy simple enough to apply and rich enough to usefully describe the incidents likely to be encountered. Figure 1 is a sample set of intrusion categories [1]:

The intrusion "names" given in Figure 1 can easily be replaced with terms tailored to the individual organization.

The focus of this exercise is to count the number of times defensive measures fail, and how badly. A common classification system (emphasis on *common*) is a prerequisite if incident responders are to communicate their true security posture. In Figure 1, they know that if they're facing a Breach 2, they need to implement containment faster than if confronting

Name	Description
Cat 6	Intruder conducted reconnaissance against asset with access to sensitive data.
Cat 3	Intruder tried to exploit asset with access to sensitive data, but failed.
Cat 2	Intruder compromised asset with access to sensitive data but did not obtain root- or administrator-level access.
Cat 1	Intruder compromised asset with ready access to sensitive data.
Breach 3	Intruder established command-and-control channel from asset with ready access to sensitive data.
Breach 2	Intruder exfiltrated nonsensitive data or data that will facilitate access to sensitive data.
Breach 1	Intruder exfiltrated sensitive data or is suspected of exfiltrating sensitive data based on volume, etc.
Crisis 3	Intruder publicized stolen data online or via mainstream media.
Crisis 2	Data loss prompted government or regulatory investigation with fines or other legal consequences.
Crisis 1	Data loss resulted in physical harm or loss of life.

Figure 1: Example of intrusion categories

a Cat 2. Why? Because in the case of a Breach 2, data theft is imminent, whereas a Cat 2 has not yet escalated to the same likelihood of negative consequences. This is what we mean by metrics that deliver decision support.

2. **Measuring the time elapsed from the moment of detection to the moment of risk reduction:** When first reading this output metric, you might ask, "What is risk reduction?" Professional incident responders speak in terms of "containment"—actions taken to remove an intruder's ability to communicate with a compromised resource (i.e., if a compromised computer is "contained," then the intruder can neither steal data from it nor issue remote commands to alter its state). Rogue, independent code that is already on the machine, however, can still take actions whether or not it is cut off from the outside world, such as to scramble or delete data—that risk remains even if the risk of data exfiltration from the compromised system is eliminated when that system can no longer contact its home base (or be contacted by it).

You may ask, "Why measure from the moment of detection to the moment of risk reduction (i.e., containment)? Why not measure time elapsed from the moment of compromise to the moment the resource is returned to a trustworthy state (i.e., recovery)?" Organizations beginning to measure output metrics should start with the more achievable goal of measuring detection-to-containment. On the "left" side of the time horizon, determining when an intrusion first occurred can be a daunting task. In Mandiant's experience, many serious intrusion victims wait months before learning they have been compromised: 243 was the median number of days from compromise to detection, and $^2/_3$ of the time a third party notified the victim of the intrusion. (Both Verizon's Data Breach Investigations Report and the Index of Cyber Security have found similar numbers.) On the "right" side of the time horizon, moving from containment to recovery is not necessarily the responsibility of the security team; usually the IT team is tasked to rebuild intrusion victims from gold master builds. Measuring that process is outside the security team's purview and can be unnecessarily demoralizing if recovery is a lengthy process.

Or you may ask, "Why does time matter at all? Shouldn't the severity of the intrusion be the most important metric?" Even if it's true that severity is ultimately the most important metric, an incident response team will rarely recognize the severity of an intrusion at the moment of first detection. At best, the team can decide whether the intrusion merits a "fast path" or a "slow path" response process. Using threat intelligence, the most advanced incident response teams identify perpetrators with a history of inflicting the most damage, or having the intention and/or capability of inflicting the most damage. Upon finding these critical threat groups active in the enterprise, the IR team responds using a "fast path," taking actions to quickly implement containment. Other intruders deserve a "slow path." Triage is not just for hospital emergency rooms.

If time is important, what is the reward for being fast? Again, using Mandiant's experience, critical threat groups do not act "at network speed" or "at the speed of light" as is often heard when speaking to government and defense officials. Rather, most data thieves need time to move beyond their initial foothold. They need to find the data they want, figure out a way to remove it, and only then exfiltrate that information. This process may take days or weeks, not nanoseconds. All the while, they must remain unseen by the victim organization and any third parties with extraordinary detection mechanisms. If at any point the defenders detect and interrupt the intruders before they can steal data, the victim organization "wins."

One public example of the role of time comes from the 2012 hack of the State of South Carolina Department of Revenue. As shown in Figure 2, it took the criminal group exactly one month to accomplish its goal [1]. Although the threat actor compromised the agency on August 13, 2012, they did not remove any data until September 13, 2012.

Recalling that Availability = MTBF / MTBF+MTTR , where MTBF is "Mean Time Between Failures" and MTTR is "Mean Time To Repair," you can see that 100% availability comes from either making MTBF infinite or making MTTR zero. In the spirit of supporting your decisions, it's clear that there comes a point where further investment in intrusion avoidance is diseconomic. From that point on, if not sooner, your investments should go towards reducing the duration of compromise, hence the very metric we suggest.

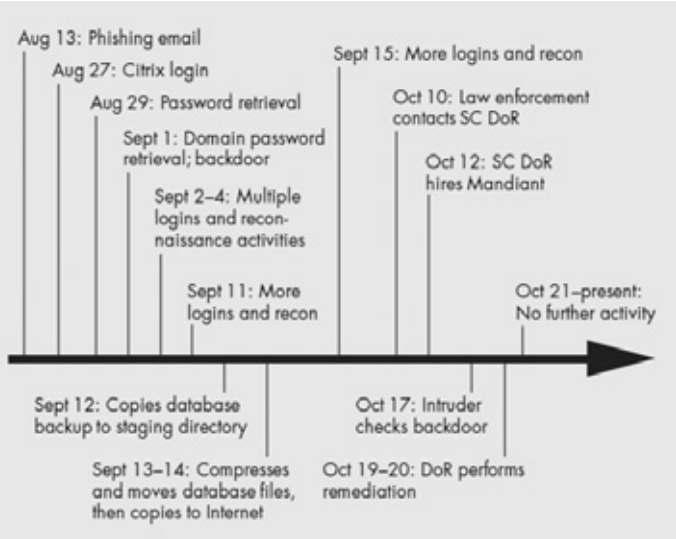

Figure 2: Timeline of a compromise

While we are at it, if there is no such thing as a "good intrusion" nor is perfect knowledge ever possible, then measuring time is the most concrete way to evaluate incident handling maturity, especially in large organizations suffering many intrusions. Beyond counting/classifying and measuring time as we argue above, we offer one simple indicator that an incident response capability is moving in the right direction: is your organization a member of FIRST? FIRST is the Forum of Incident Response and Security Teams, a global security group founded in 1989 and consisting of nearly 300 members (Figure 3).

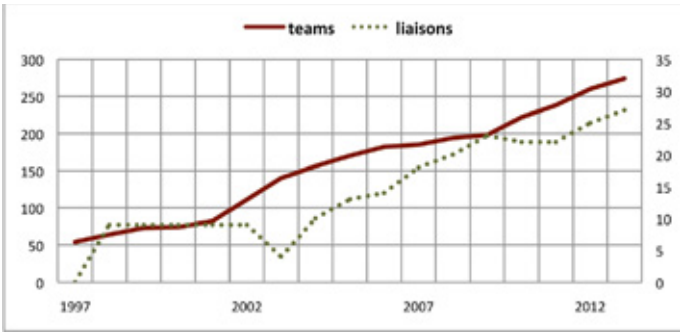

Figure 3: FIRST membership by year

FIRST membership demonstrates a commitment to creating and maintaining a formal incident detection and response program, backed by an audit of a candidate member's capability and recommendation by two current FIRST member teams. By successfully joining FIRST, an organization says, "We are committed to incident response as a core element of a security program." The growth in FIRST membership since 1989 parallels the rise of incident response as a viable security strategy, complementing (and some might say now, partially displacing) incident avoidance.

Perhaps we should have admitted it at the outset, but we are pessimists who prefer to think of ourselves as realists. If you are a security program manager, then prepare for when you get hacked next. If you are a statistician, then get the data—you can always throw it away later.

References

[1] Taken from Richard Bejtlich, *The Practice of Network Security Monitoring* (No Starch Press, 2013).