Paxson
Spring 2017

CS 161
Computer Security

Discussion 12

Week of April 24, 2017

Question 1 ***Detection strategies*** **(20 min)**

Suppose you are responsible for detecting attacks on the UC Berkeley network, and can employ host-based monitoring (a HIDS) that can inspect the keystrokes that users enter during their shell sessions. One particular attack you are concerned with is malicious modification or deletion of files in the directory `/usr/oski/config/`.

(a) One method of detection is called "signature matching." This involves looking for particular well-defined patterns in traffic that are known to represent malicious activity. Give a couple of examples of signatures you can use to detect these attacks. What are some limitations of this approach?

(b) Another approach is to search for behaviors. Instead of looking for known attacks, the detector might use knowledge of the system to look for suspicious sets of actions. Give two examples of host-based behavioral detection. Be specific as to how your examples differ from signature matching that looks for known attacks. What are some problems with this approach?

(c) Suppose now we aim to detect modifications to *any* files in `/usr/oski/config/` using the following procedure. Each night, we run a cron job that checksums all of the files in the directory using a cryptographically strong hash like SHA256. We then compare the hashes against the previously stored ones and alert on any differences. (This scheme is known as "Tripwire.")

Discuss issues with false positives and false negatives.

(d) Continuing the previous scenario, suppose the attacker was able to subvert the operating system. Can you think of a procedure (which might be expensive in terms of labor) by which an operator could still detect the modified files?

Question 2 ***Detecting Web Attacks*** **(15 min)**

At this year's annual *Grasses For The Masses* home & garden convention, in beautiful Fairfax California, the startup *Lazer Lawns*—which specializes in producing so-juicy-looking-you-just-wanna-eat-it artificial turf—experienced a live SQL injection attack from the audience while showcasing their new high-end collection of silver-ionized heat-repellent blades—what a disaster! After firing the organizer of the event and hiring a CS161-educated security expert, Grasses For The Masses now plans to install a NIDS that watches the free WiFi next year. Moreover, Lazer Lawns has learned the hard way to make sure they have a HIDS[1] protecting their assets.

[1] Given they have to demo their software in many different environments, they've learned that they shouldn't rely on being able to employ a NIDS, hence their emphasis on using a HIDS.

As a potential advisor to either Grasses For The Masses or Lazer Lawns, consider the some prevalent web attacks: XSS (both reflected and stored) and SQL injection.

(a) For each attack, devise one or more concrete strategies based on signature, behavioral, anomaly, or specification-based detection. Include a discussion of false positives and false negatives.

(b) Explain whether a network-based or host-based deployment approach makes more sense for your devised detection strategy (or if it doesn't really matter). Does the deployment angle have an effect on your detection rates?

Question 3 ***Detection Tradeoffs*** **(15 min)**

Suppose that S is a network-based intrusion detector that works by passively analyzing individual UDP and TCP packets. Suppose that A is a host-based intrusion detector that is a component of the browser that processes and analyzes individual URLs before they are loaded by the browser. Suppose S has false positive rate S_P and false negative rate S_N, and A has false positive rate A_P and false negative rate A_N.

Your company decides to build a hybrid scheme for detecting malicious URLs. The hybrid scheme works by combining scheme S and scheme A, running both in parallel on the same traffic. The combination could be done in one of two ways. Scheme H_E would generate an alert if for a given network connection either scheme S or scheme A generates an alert. Scheme H_B would generate an alert if both scheme S and scheme A generate an alert for the same connection. (Assume that there is only one URL in each network connection.)

(a) Assuming that decisions made by S and A are well-modeled as independent processes, and ignoring any concerns regarding evasion, what can you say about the false positives and false negatives of H_B and H_E? In terms of S_P, S_N, A_P, A_N, what are the false positive and false negative rates for H_B and H_E?

(b) If deploying the hybrid scheme in a new environment, is one of H_E and H_B clearly better? If so, which one, and why? If not, what environment parameters would help determine whether H_E or H_B is better?