Tutorial #10

Security in Computing COSC2356/2357

Q1: Why intrusion detection system (IDS) is important?

Answer:

The purpose of intrusion detection system (IDS) is to detect attacks before, during, and after they occur. The basic approach employed by IDS is to look for "unusual" activity. Once an intrusion is detected, we want to respond to it. In some cases, we obtain specific information and a reasonable response is fairly obvious. For example, we might detect a password guessing attack aimed at a specific account, in which case we could respond by locking the account.

In Japan, the total number of crimes involving the Internet was almost 60% higher in the first half of 2001 than in 2000? Furthermore, Internet fraud has increased by 94%! According to a report mentioned in the Computer Crime and Intellectual Property section, in from 1995 – 2000, Hong Kong witnessed 26 times increase in cybercrime cases. In 1995, the number of crimes totalled to 14 and in 2000, the crimes had increased to 368. The most popular social media networks see as many as 20 million probes per week! These probes are generally caused due to reasons, such as unauthorized access gained by attackers, malware and authorized users try to get additional privileges. The answer to this is an *intrusion detection system*.

Intrusion detection system is like a burglar alarm for a computer network. It monitors the flow of traffic and facilitates information systems to deal with the attacks. This system protects an enterprise setup by identifying, logging, reporting and sending alarm whenever there is an abnormality. In conclusion, the main task of intrusion detection system is to detect attacks and possibly repel them.

Q2: The anomaly-based intrusion detection example presented in Lecture-10 is based on file-use statistics.

a) Many other statistics could be used as part of anomaly-based IDS. For example, network usage would be a sensible statistic to consider. List five other statistics that could reasonably be used in anomaly-based IDS.

b) Why might it be a good idea to combine several statistics rather than relying on just a few?

c) Why might it not be a good idea to combine several statistics rather than relying on just a few?

Answer:

a) Commands issued, typing speed, mouse movements, activity versus inactivity, time of use, among many, many other possibilities.

b) More statistics would give a clearer view of the user's activity.

c) More statistics would make it slower, and it might also tend to give more false alarms, since a legitimate user is more likely to vary from one of the stats.

Q3: Suppose in a host-based anomaly detection system, over an extended period of time, Alice has accessed four files, F_0, F_1, F_2, F_3, at the rates H_0, H_1, H_2, H_3, respectively, where the observed values of the H_i, for i = 0,1,2,3, *are* given in Table 3.1.

Table 3.1: Alice's Initial File Access Rates

H_0	H_1	H_2	H_3
0.20	**0.10**	**0.05**	**0.20**

Now suppose that, over a recent time interval, Alice has accessed file F_i at the rate A_i, for i = 0,1,2,3, as given in Table 2.2.

Table 3.2: Alice's Recent File Access Rates

A_0	A_1	A_2	A_3
0.10	**0.05**	**0.15**	**0.25**

Given, the statistics (*S*) of comparison of long-term access rates (*H_i*) to the current rates (*A_i*) is considered as normal if *S* < 0.1. Find the answers for the following questions:

a) Do Alice's recent file access rates represent normal use?

b) Suppose the previous values of access rates are weighted at 80%, while the current values are weighted 20%. What are Alice's updated file access rates?

Answer:

(a) The statistics (***S***) of comparison of **long-term access rates** (***H_i***) to the current rates (***A_i***) is computed as follows:

$$S = (H_0 - A_0)^2 + (H_1 - A_1)^2 + (H_2 - A_2)^2 + (H_3 - A_3)^2$$

$$= (0.20 - 0.10)^2 + (0.10 - 0.05)^2 + (0.05 - 0.15)^2 + (0.20 - 0.25)^2$$

$$= 0.025$$

From the value of **S = 0.025** which is less than **0.1**, we conclude that Alice's recent use is normal.

(b) To update the historical access rates, we use a moving average that combines the previous values with the recently observed rates. Given that the previous values are weighted at 80%, while the current values are weighted 20%.

Update Alice's long-term history values ***H_i*** according to the formula:

$$H_i = 0.2 * A_i + 0.8 * H_i \text{ (for, i = 0,1,2,3.)}$$

$$H_0 = 0.2 * 0.10 + 0.8 * 0.20 = 0.18$$

$$H_1 = 0.2 * 0.05 + 0.8 * 0.10 = 0.09$$

$$H_2 = 0.2 * 0.15 + 0.8 * 0.05 = 0.07$$

$$H_3 = 0.2 * 0.25 + 0.8 * 0.20 = 0.21$$

Therefore, Alice’s updated File Access Rates are as follows:

Table: Alice's Updated File Access Rates

H_0	H_1	H_2	H_3
0.18	0.09	0.07	0.21

Q4 Recall that the anomaly-based IDS example presented earlier is based on file-use statistics. The expected file use percentages are periodically updated using an equation, which can be viewed as a moving average.

a) Why is it necessary to update the expected file use percentages?
b) When we update the expected file use percentages, it creates a potential avenue of attack for Trudy. How and why is the case?

Answer:

a) Use changes over time, so if these values do not change, you will soon get many false alarms.
b) Trudy can simply “go slow” and eventually convince the IDS that her actions are normal.

Q5 Recall the concept of **Proof-of-Work (PoW)** that was discussed in Lecture 7 and Tutorial 7. Think and discuss how **PoW** can be used to protect email spamming.

Answer:

The email spamming can be protected by using the concept of Proof-of-Work (PoW). For example, the user’s email client can be forced to generate the hash of the email with four leading zero and to send the hash along with email content.