Homework lecture 2
Prevention-and-detection

Harald Vranken
harald.vranken@ou.nl
February 10, 2020

Question 1
IPsec operates at the network layer. Using IPsec in tunnel mode allows to set up a VPN between two end-points. The payload of an IPsec datagram is secured using cryptography by either encrypting the payload or adding authentication headers. Is such a VPN compatible with NAT? Explain why, or why not, this is the case. And if not, how could this be resolved?
(Note: we did not cover NAT in the lecture, but I presume you are familiar with the basic concept of network address translation.)

Answer: NAT translates the private IP address of an internal host into the user's public IP address. The NAT device also changes the port, to map the IP address and port of the internal host to the public IP address and port. Hence, the NAT device uses the public port to actually identify the internal host.
In IPsec tunnel mode, a separate IP header is added to the network layer datagram. This IP header contains IP addresses of source and destination, and NAT is able to modify these. The ports however are included in the transport layer header, which is part of the payload of an IPsec datagram. In case of AH, the payload is equipped with authentication headers. Modifying a port will break this authentication. In case of ESP, the payload is encrypted and hence the port cannot be identified nor modified.
This can be resolved by using NAT-T (NAT traversal) using IPsec over UDP, where the IPsec datagram is encapsulated in an UDP-segment. The NAT device can change the network and transport layer headers of the UDP message, and hence there is no need to touch the IPsec datagram.
An alternative solution is to let the NAT device be the end-point of the IPsec tunnel. The NAT device is then able to modify both IP addresses and port numbers before (when sending) or after (when receiving) applying IPsec to a packet. Note however that this breaks end-to-end security, since the tunnel is now between the end-point in the external network and the NAT device, and not between the end-point in the external network and the end-point in the internal network.
Port forwarding is a common technique to traverse a NAT device, for instance for a server in the internal network that should be reachable form the external network. The NAT device then configures a port on its public interface, and forwards traffic that is directed to this port, to the server in the internal network. This however is not possible when applying IPsec in tunnel mode, since the NAT device cannot see the port number at the transport layer.

Question 2
When interpreting the alarms generated by an IDS, you should be aware of the base-rate fallacy. In the lecture we looked at an example (see slide 34) of an IDS with a false alarm rate of 99%. What should the accuracy of this IDS be to reduce the false alarm rate below 50%?

Answer:
Let the IDS accuracy be *x* (where $0 \leq x \leq 1$).
Of the 100 malicious events, there will be 100*x* true positives, and 100(1-*x*) false negatives.
Of the 1,000,000 benign events, there will be 1,000,000*x* true negatives, and 1,000,000(1-*x*) false positives.
Hence, the number of positives reported is 100*x* + 1,000,000(1-*x*), of which the 100*x* true positives should be at least 50%.
Solving $100x / (100x + 1{,}000{,}000(1-x)) \geq 0.5$ yields $x \geq 10{,}000/10{,}001 \approx 0.9999$.
The accuracy should therefore be larger than 0.9999 (ie. 99.99%).

Question 3
Flow entries are stored in the flow cache. Flow entries expire on the occurrence of certain events, such as observing an TCP packet with a FIN or RST flag that indicates the end of the flow. Flow entries can also expire on timeouts, which can be either active timeouts or idle timeouts. What is the difference between active and idle timeouts? And what can be the impact when configuring too low or too high timeout values?

Answer: Active timeouts are used to force long-living flows to expire. Typical active timeout values range from 2 to 30 minutes. Since a flow is not actually ended on an active timeout, the entry is not removed from the cache, but counters are reset and start and end times are updated.
Idle timeouts are used to consider inactive flows, for which no packets have been observed for a specified amount of time, as ended. Typical timeout values range between 15 seconds to 5 minutes.
The configured timeout values have impact on the total number of flows exported and the number of flow entries stored in the flow cache. Longer timeout values result in higher aggregation of packages into flow records, which implies that flows reside longer in the cache resulting in higher cache utilization (ie. more successful cache entry lookups). This is generally positive to reduce the load on flow collectors, but implies that it takes longer before flows become visible to the data analysis stage.
Reducing timeout values reduces cache utilization and results in more flow records, which can cause capacity problems as the flow cache gets overloaded. At that moment, emergency expiration kicks in to expire cache entries prematurely to free up the flow cache and allow new entries to be inserted. This however causes more flow records, and flow data that is not expired consistently, which may impact subsequent data analysis.
Hence, timeout values should not be too high to allow timely analysis, but also not too low to prevent cache overflows and inaccurate flow data.

Question 4
SSH attacks and DDoS attacks can be detected successfully with a flow-based IDS. What flow metrics are used best to detect these attacks?

Answer: The number of packets-per-flow (ppf) is the best metric to detect these attacks. SSH attacks typically show a low number of ppf in the scan phase and a significantly larger, nearly constant number of ppf in the brute-force phase.

DDoS attacks are typically detected by counting the number of ppf per second from the same IP address that contain only few packets.