CS.201P: Introduction to Computer Security 2/15/2019

Homework 3

Due Thursday 2/28/2019, 11:59pm, by upload to gradescope.com

Solve the following problems from the [GT] textbook:

1. Exercise R-5.3 & R-5.5

R-3 Can two network interfaces have the same MAC address? Why or why not?

General not. Network interfaces are typically identified by a hardware-specific identifier known as its media access control address (MAC address). A MAC address is a 48-bit identifier represented by six pairs of hexadecimal digits; the first 24 bits show the organization that issued the MAC address, ie IEEE, and sometimes is used to identify a particular brand or model of the network interface; the last 24 bits give the manufacturer a basically unique MAC address for different model instances, which have a total of 2^24 possibilities. MAC addresses are used in the link layer to identify the devices in a network and each MAC address will only appear once in a switch's MAC address table. If a there are more than one NICs on the same LAN that share the same MAC address, it will cause all sorts of communication issues, like MAC flapping.

But it is possible. If the manufacturer may reuse the MAC address, the probability that the manufacturer's two the devices have the identical MACs in a network is also small, about one in a million. And it is possible via MAC spoofing. If you are seeing the same MAC address on multiple ports then it is possible that someone is in your network doing some not nice things.

R-5 Can two network interfaces have the same IP address? Why or why not?

It is possible to have two network interfaces with the same IP address in Network Address Translation (NAT). These addresses would not be exposed on the general outside Internet while sharing a public IP address. They would exist inside two separate LANs that use NAT internally to route packets within the LAN from addresses in the general Internet. Absent NAT, IP addresses on the general Internet are unique; otherwise, routing would not make sense.

2. Exercise R-5.6

R-6 Show why installing static ARP tables on the machines of a local-area network does not prevent a malicious machine from intercepting traffic not intended for it.

Static ARP tables, requires a network administrator to manually specify a router's ARP cache to assign certain MAC addresses to specific IP addresses. When using static ARP tables, ARP requests to adjust the cache are ignored, so ARP spoofing of that router is impossible. This requires the inconvenience of having to manually add entries for each device on the network, however, and reduces flexibility when a new device joins the network, but significantly mitigates the risk of ARP cache poisoning.

This solution does not prevent an attacker from spoofing a MAC address to intercept traffic intended for another host on the network. An ARP request for an IP address is sent to all machines on the local area network that request for an IP address, there is no authentication scheme so anyone can claim to have the requested address. Installing an ARP tables on the machine of a local area network does not prevent another machine from looking through the data since there's no authentication so the data that was meant for someone else can be caught by someone else.

3. Exercise R-5.15

R-15 How is it that a machine of a private network behind a NAT router can make a connection with a web server on the public Internet?

To translate between private and public IP addresses, the NAT router maintains a lookup table that contains entries of the following form:
(private source IP, private source port, destination IP, public source port)

A NAT router dynamically rewrites the headers of all inbound and out- bound TCP and UDP packets as follows. When a machine on the internal network attempts to send a packet to an external IP address, the NAT router creates a new entry in the lookup table associated with the source machine's private IP address and the internal source port of the transmitted packet. Next, it rewrites the source IP address to be that of the NAT device's public IP, opens a new public source port, and rewrites the IP header's source port field to contain the newly opened port. This public port and the destination IP address are recorded alongside the private source IP and private internal port in the NAT device's lookup table. The NAT device also adjusts any checksums contained in the packet, including those used by IP and TCP/UDP, to reflect the changes made. The packet is then forwarded to its destination.

On receiving a response, the NAT router checks its lookup table for any entries whose public source port corresponds to the destination port of the inbound packet and whose

destination IP address (recorded because of the previous outbound packet) corresponds to the source IP of the inbound packet. Finally, the NAT router rewrites the IP headers of the inbound packet according to the lookup table, so that the packet is forwarded to the correct private IP address and private port.

This process effectively manages outbound traffic, but places several restrictions on the possibilities for inbound traffic. An external machine has no way of initiating a connection with a machine on the private network, since the internal machine does not have a publicly accessible IP address. This can actually be seen as a security feature, since no inbound traffic from the Internet can reach the internal network. Thus, in many ways, NAT devices can function as firewalls, blocking risky contact from the external Internet.

Network Address Translation is not a perfect solution. In fact, it violates the ideal goal of end-to-end connectivity for machines on the Internet by not allowing direct communication between internal and external parties. In addition, NAT may cause problems when using several protocols, especially those using something other than TCP or UDP as a transport-layer protocol. Still, NAT has been crucial in delaying the exhaustion of the IPv4 address space and simplifying home networking.

4. Exercise C-5.3

C-5.3 Show how to extend the man-in-the-middle attack described in Section 5.2.3 to intercept all documents sent to a printer in a local-area network.

Since ARP protocol lacks an authentication scheme, any computer on the network could claim to have the requested IP address. In fact, any machine that receives an ARP reply, even if it was not preceded by a request, will automatically update its ARP cache with the new association. Because of this shortcoming, it is possible for malicious parties on a LAN to perform the ARP spoofing attack.

This attack is relatively straightforward. An attacker, Eve, simply sends an ARP reply to a target, who we will call Alice, who associates the IP address of the LAN gateway, who we will call Bob, with Eve's MAC address. Eve also sends an ARP reply to Bob associating Alice's IP address with Eve's MAC address. After this ARP cache poisoning has taken place, Bob thinks Alice's IP address is associated with Eve's MAC address and Alice thinks Bob's IP address is associated with Eve's MAC address. Thus, all traffic between Alice and Bob (who is the gateway to the Internet) is routed through Eve, as in the following Figure 8 (from textbook).

Here, Printer is Alice, and Bob is still the gate to internet. The attack, Eve, in the same LAN, build a connection between Alice and Bob, where Alice regards Eve is Bob and Bob regards Eve as Alice, thus can obtains any fill that is sent to the printer.

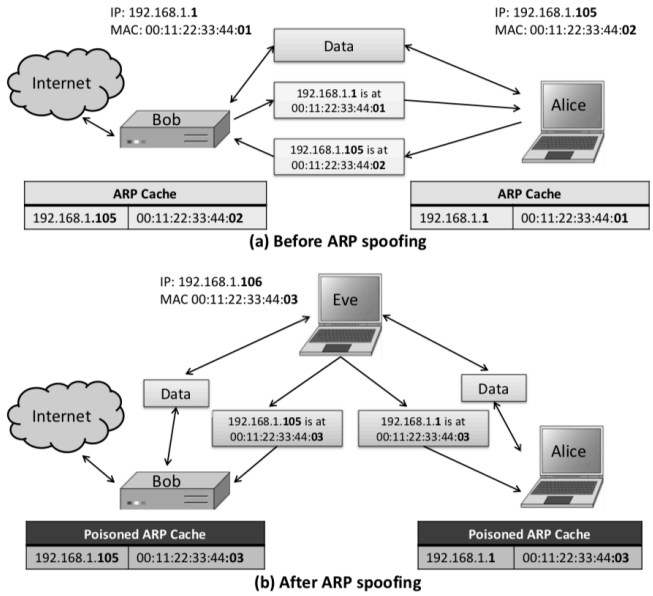


Figure 8: ARP spoofing enables a man-in-the-middle attack: (a) Before the ARP spoofing attack. (b) After the attack.

5. Exercise C-5.4

C-5.4 Suppose you suspect that your session with a server has been intercepted in a man-in-the-middle attack. You have a key, K, that you think you share with the server, but you might be only sharing it with an attacker. But the server also has a public key, K_P, which is widely known, and a private secret key, K_S, that goes with it. Describe how you can either confirm you share K with the server or discover that you share it only with a man-in-the-middle. Also, be sure your solution will not be discovered by a packet sniffer.

The server has a public key Kp and a private key Ks. If there is a man-in-the-middle, then it doesn't know what is the private key of server. We have to utilize this observation.

So we ask server to encrypt our key K with its private key Ks and send back to us. We then use server's public key to decrypt the data. If we can use our key K to decrypt the content message, then we know it is real server. The man-in-the-middle can't have server's private key and the message from hacker can't be decrypted in meaningful message, and it works even if it is sniffed.

6. Exercise R-5.11 & C-5.7

R-11 Describe how sequence numbers are used in the TCP protocol. Why should the initial sequence numbers in the TCP handshake be randomly generated?

TCP ensures reliable transmission by using a sequence number that is initialized during the three-way handshake. Each subsequent transmission features an incremented sequence number, so that each party is aware when packets arrive out of order or not at all.
First, a client sends a packet to the desired destination with the SYN flag (short for "synchronization") set. This packet includes a random initialization for a sequence number, which is used to ensure reliable ordering of future data transmissions.

In response, the server replies with a packet marked with both the SYN and ACK (short for "acknowledgment") flags, known as a SYN-ACK packet, indicating that the server wishes to accept the connection. This packet includes an acknowledgment number, which is set to one more than the received sequence number, and a new random sequence number. On sliding window protocol, when sending data, the sender checks the sequence number of the packet to be sent, and only continues sending if this number is less than the last acknowledgment number plus the current size of the receive window.

Finally, on cumulative acknowledgment scheme, the client responds with an ACK packet to indicate a successful connection has been established. The final ACK packet features an acknowledgment number set to one more than the most recently received sequence number, and the sequence number set to the recently received acknowledgment number.

The figure is as below (from textbook). Generally, the sequence numbers are used for ensure the order and quality of data transfer.

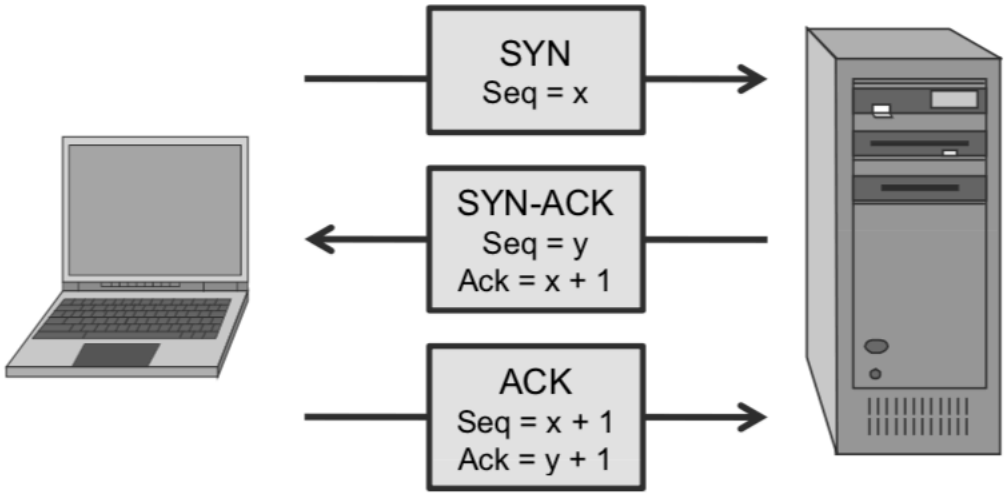
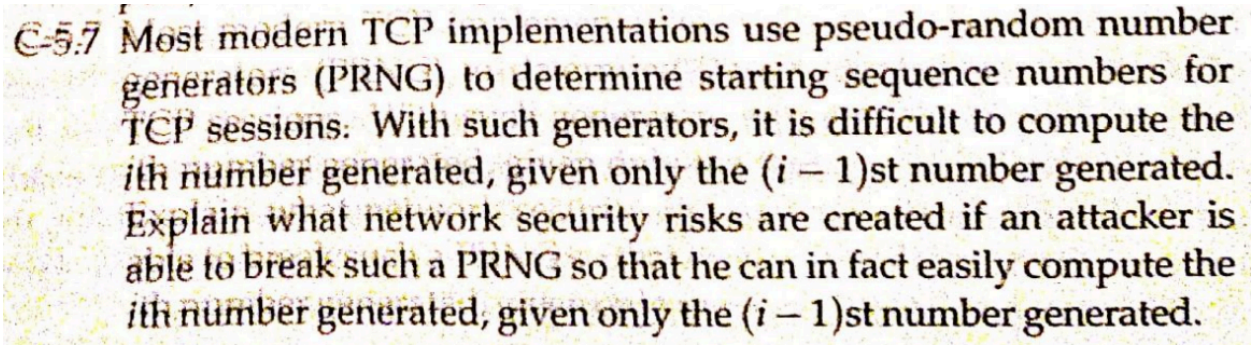


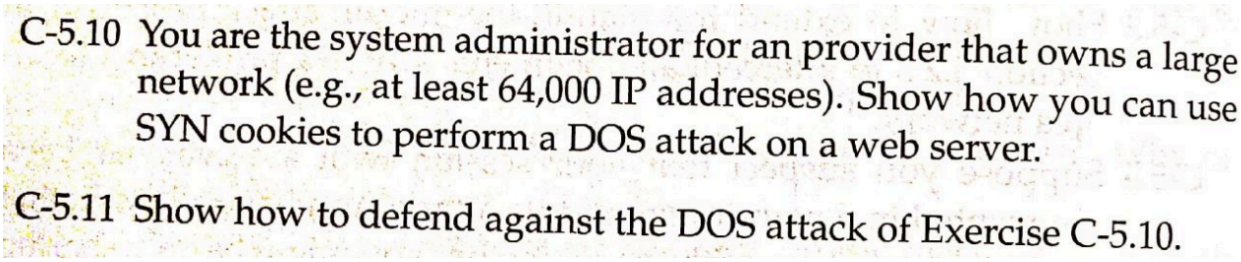
Figure 15: The three-way TCP handshake.

The number choices are meant to defeat attacks against TCP based on predicting initial sequence numbers. Early TCP stacks implemented sequence numbers by using a simple counter that was incremented by 1 with each transmission. Without using any randomness, it was trivial to predict the next sequence number, which is the key to TCP session hijacking attack, which might later cause flood packets on network or let the attack obtain secret information or generate fake data.


C-5.7 Most modern TCP implementations use pseudo-random number generators (PRNG) to determine starting sequence numbers for TCP sessions. With such generators, it is difficult to compute the ith number generated, given only the $(i-1)$st number generated. Explain what network security risks are created if an attacker is able to break such a PRNG so that he can in fact easily compute the ith number generated, given only the $(i-1)$st number generated.

As discussed above, there may come some kind of TCP session hijacking attack. If the attacker creates a new spoofed TCP session, and with combination with other network attacks, it can then may be possible to inject a packet containing a command that creates a connection back to the attacker, cause an ACK storm. may create a man-in-the-middle situation, might later cause flood packets on network or let the attack obtain secret information or generate fake data.

7. Exercise C-5.10 & C-5.11


C-5.10 You are the system administrator for an provider that owns a large network (e.g., at least 64,000 IP addresses). Show how you can use SYN cookies to perform a DOS attack on a web server.

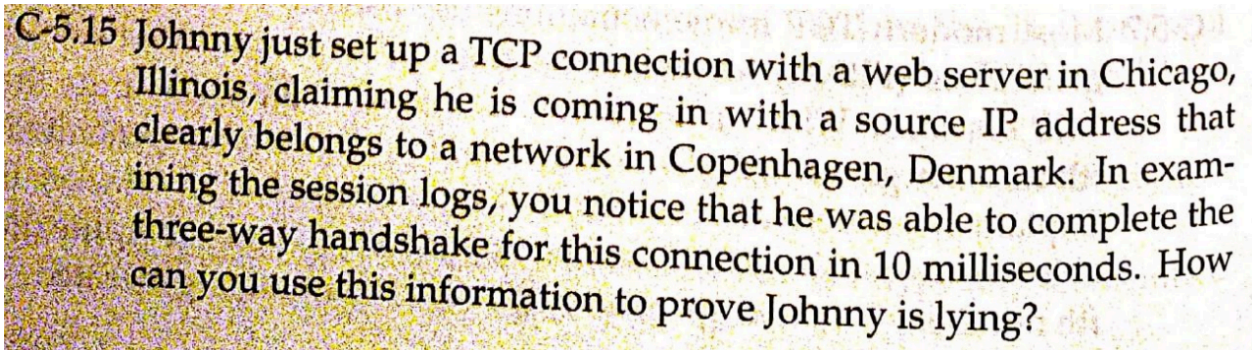
C-5.11 Show how to defend against the DOS attack of Exercise C-5.10.

SYN cookies are a way to mitigate SYN flood attacks, thus causing a DOS on a web server. Since as the system administrator, we can create a SYN cookie: when the server receives an "ACK" packet from the client, it always reports a mismatch or does not send the final ACK. Ignore Server and client then go through the "SYN + ACK" and "ACK" packets without having to connect successfully. It can gradually fill the packet queue, run out of the resources or bandwidth and cause DOS across the network over the network.

The solution can be more effectively manage half-opened connections, including implementing a special queue for half-open connections and not allocating resources for a

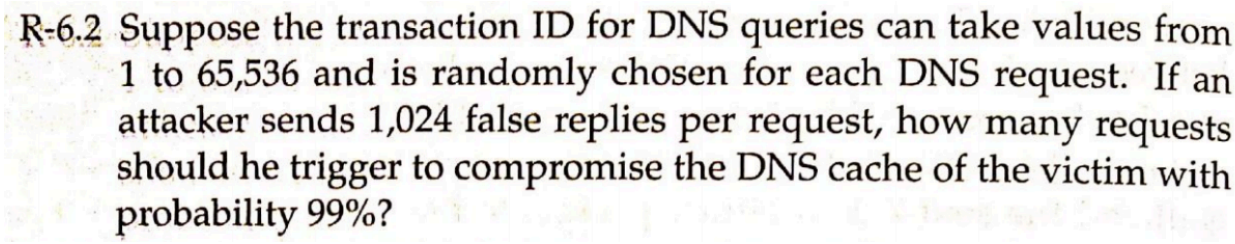
TCP connection until an ACK packet has been received. Or easily limit the number of connections for IP addresses within a specific range, set a shorter session time limit for the client to connect, and even temporarily disable a specific range of IP addresses.

8. Exercise C-5.15


C-5.15 Johnny just set up a TCP connection with a web server in Chicago, Illinois, claiming he is coming in with a source IP address that clearly belongs to a network in Copenhagen, Denmark. In examining the session logs, you notice that he was able to complete the three-way handshake for this connection in 10 milliseconds. How can you use this information to prove Johnny is lying?

The distance from Chicago to Denmark is about 6,839 kilometers. The speed of light is 299.792 km / s. It takes 0.023 seconds without delay and re-transmition to find a method. Since the TCP handshake is three-way, it needs at least 0.023 * 3 = 0.069 seconds, which is more than 10 million seconds. So we can tell Johnny that he is lying.

9. Exercise R-6.2


R-6.2 Suppose the transaction ID for DNS queries can take values from 1 to 65,536 and is randomly chosen for each DNS request. If an attacker sends 1,024 false replies per request, how many requests should he trigger to compromise the DNS cache of the victim with probability 99%?

1) Since per request the attacker send 1024 request in 65536 ID, so the chance of success is p = 1024/65536 = 0.015625 per time;
2) So the failure rate of one time is q = 1- p = 0.984375. With k times try, the failure rate of not success even one time is Q = q^k;
3) Therefore, to achieve a success rate of more than 99%, we should try k times with:
   Q <= 1- 99% = 1%,
   100 <= (1/q)^k
   k >= ln(100)/ln(1/q) = 292.4
   so attacker must more than 293 times.

10. Exercise R-6.7 & R-6.8

R-6.7 Explain how a stateless firewall would block all incoming and outgoing HTTP requests.

R-6.8 How can SSH be used to bypass firewall policy? What can a network administrator do to prevent this circumvention?

A stateless firewall can block all inbound and outbound packets on port 80 or IP. The farewall policy can only manage source / destination IPs, ports, and so on. If the policy is based on the source IP, the source IP packets in the blacklist are always blocked. A similar situation applies to the destination IP in the blacklist. For incoming and outgoing messages, the handshake can be successfully set. This means that all requests are blocked.

SSH can be used to bypass firewall policies by entering a tunnel that encapsulates all traffic and makes packets look like normal traffic. To prevent this, the network administrator must perform a reverse SSH tunnel.

11. Exercise R-6.14

R-6.14 What are the main differences between WEP and WPA? What are the different possible modes under the WPA standard?

WEP encrypts each data frame using RC4 stream cipher. Its security key intense is mainly in 128 bits. It draws back WEP key recovery, unauthorized decryption, violation of data integrity, and poor key management.

Once the weaknesses in RC4 and WEP were published, Wi- Fi Protected Access (WPA), a more complex authentication scheme is developed. Its security key intense is mainly in 256 bits. It implementes in integrity checks and provides features like Message integrity and Temporal Key Integrity Protocol (TKIP).

12. Exercise R-6.15

R-6.15 Explain why deep packet inspection cannot be performed on protocols such as SSL and SSH.

Protocols such as SSL or SSH have encypted data into very long bits data, and the encrypted payloads of the packets are encrypted with a secret key shared by the client and server but not known to the firewall.

13. Exercise C-6.3

C-6.3 Suppose Alice sends packets to Bob using TCP over IPsec. If the TCP acknowledgment from Bob is lost, then the TCP sender at Alice's side will assume the corresponding data packet was lost, and thus retransmit the packet. Will the retransmitted TCP packet be regarded as a replay packet by IPsec at Bob's side and be discarded? Explain your answer.

No. IPsec treats a retransmitted TCP packet as a new IPsec packet. It is depended on TCP to notice the packet is duplicate.

14. Exercise C-6.4

C-6.4 An alternative type of port scan is the ***ACK scan***. An ACK scan does not provide information about whether a target machine's ports are open or closed, but rather whether or not access to those ports is being blocked by a firewall. Although most firewalls block SYN packets from unknown sources, many allow ACK packets through. To perform an ACK scan, the party performing the scan sends an ACK packet to each port on the target machine. If there is no response or an ICMP "destination unreachable" packet is received as a response, then the port is blocked by a firewall. If the scanned port replies with a RST packet (the default response when an unsolicited ACK packet is received), then the ACK packet reached its intended host, so the target port is not being filtered by a firewall. Note, however, that the port itself may be open or closed: ACK scans help map out a firewall's rulesets, but more information is needed to determine the state of the target machine's ports. Describe a set of rules that could be used by an intrusion detection system to detect an ACK scan.

A target system is presented with a packet with the ACK flag set with a sequence number of zeros to an interesting port. Since generally the sequence number is not zero, there is a violation of TCP rules associated with that parameter, and the target sends back a RST. The presence of the RST provides an attacker with a good indication that the host is alive, but behind some form of filtering. Here are two rules for detecting the behavior:

alert tcp 172.16.16.0/24 any -> 172.16.17.0/24 any (msg: “Potential Ack Scan”; flags:A; ack:0; sid: 10001;),

alert tcp 172.16.16.0/24 any -> 172.16.17.0/24 any (msg: “Ack and RST detected-Potential Ack Scan”; flags:AR; sid: 10002;)

15. Exercise C-6.11

C-6.11 Describe the types of rules that would be needed for a rule-based intrusion detection system to detect a SYN flood attack.

There are many NIDS tools such as Snort, Bro, Suricata etc. We define following rules to detect SYN flood:

- Appoint source ip/port and destination ip/port
- Log if the handshake is acked and count the number of un-acked handshake
- If many un-acked handshake requests are from same ip, we suspect there is a SYN flood attck from the ip.

16. Exercise C-6.12 & C-6.13. [As a bonus, show how to do attack in C-6.13 without storing any data. This can be done using some cryptographic tools...]

C-6.12 The *coupon collector* problem characterizes the expected number of days that it takes to get n coupons if one receives one of these coupons at random every day in the mail. This number is approximately $n \ln n$. Use this fact to compare the number of TCP connections that are initiated in a sequential port scan, going from port 1 to 65535, directed at some host, to the expected number that are requested in a random port scan, which requests a random port each time (uniformly and independently) until it has probed all of the ports.

C-6.13 Describe a modification to the random port scan, as described in the previous exercise, so that it still uses a randomly generated sequence of port numbers but will now have exactly the same number of attempted TCP connections as a sequential port scan.

There are 65535 ports in total for our problem.
A direct sequential port scan need to scan for 65536 times.

A random scan model takes: 65536 * ln 65536 = 65536 * 11.09 = 726817 It takes only 10X scan times.
We can modify our random algorithm to only pick up ports not scanned yet. It can ensure the program will always reduce our port targets in each scan iteration. The expectation of this new model will have the same number of connection as direct port scan.

17. Exercise R-7.4

R-7.4 Describe what information about a web server is stored in an SSL server certificate.

According to Web Server Certificates section, SSL contains following information:

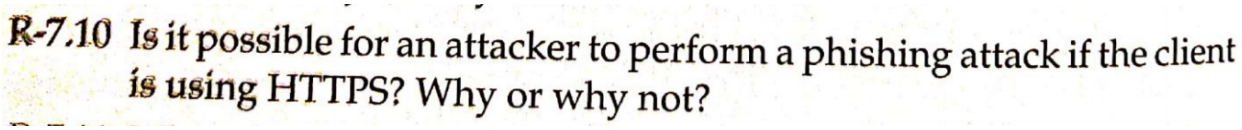
- Name of the CA
- Unique serial number
- Expiration date
- Domain name of the web site
- Organization operating the web site and its location
- Identifier of the public-key cryptosystem used by the web server (e.g.,1,024-bit RSA)
- Public key used by the web server in the HTTPS protocol
- Identifier of the cryptographic hash function and public-key cryptosystem used by the CA to sign the certificate (e.g., SHA-256 and 2,048-bit RSA)
- Digital signature over all the other fields of the certificate

18. Exercise R-7.9

R-7.9 Can a cross-site scripting attack coded in Javascript access your cookies? Why or why not?

Yes.
An attacker can use XSS to send a malicious script to an unsuspecting user. The end user's browser has no way to know that the script should not be trusted, and will execute the script. Because it thinks the script came from a trusted source, the malicious script can access any cookies, session tokens, or other sensitive information retained by the browser and used with that site. These scripts can even rewrite the content of the HTML page.
The Javascript code may not have permission to do some restricted operations, but it can create window and ask user to execute malicious code with higher execution permission. It can also ask user to load content from other website and setup a connection to third-part website. The cookies might leak to third-party website and hackers can analyze those cookies and get confidential user information.

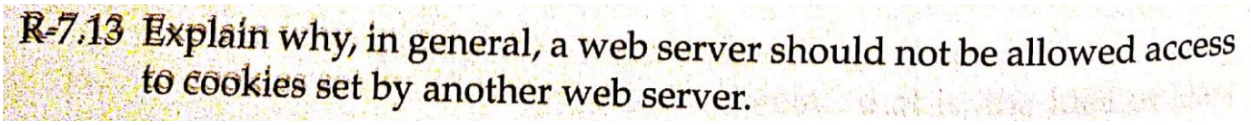
19. Exercise R-7.10


R-7.10 Is it possible for an attacker to perform a phishing attack if the client is using HTTPS? Why or why not?

Yes, it is possible that an attacker can perform phishing attack even if the client is using HTTPS. Phishing is an attempt to obtain sensitive information by masquerading as a trustworthy entity and HTTPS trust indicator could be spoofed:
An attacker can adopt man-in-the-middle attack mechanism to spoof the target user.

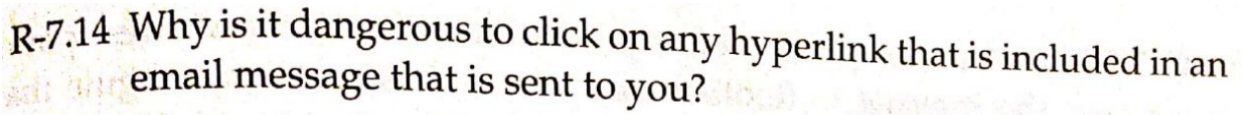
- The attacker can catch the request and forward it to the destination websites when a user tries to connect to the website,
- The attacker can monitor the communication from beginning and it will have all of the transmitted data such as keys.
- Attacker can even modify the content in transition and insert malicious code.

20. Exercise R-7.13


R-7.13 Explain why, in general, a web server should not be allowed access to cookies set by another web server.

It will be very dangerous if a web server could access cookies for another web server because sensitive information could be saved in the cookies. Cookies might be saved on disk without properly encrypted. It might log user's password, credit card information etc. If a attacker website can access to those cookies, it can get user information and might also modify the cookie. The connection state to the web server with the modified cookie could be modified as well.

21. Exercise R-7.14


R-7.14 Why is it dangerous to click on any hyperlink that is included in an email message that is sent to you?

Hyperlink might not be verified to be safe url. Sometimes, it showed a text link but its actual web page link is different from the text. It could direct user to a phishing website or download some malicious virus. The phishing website might ask for user's account information and get some sensitive information. Before clicking the website, user need to check the hyperlink by inspecting the target url and determine whether it could be trusted.

22. Exercise C-7.8. The security property your solution should provide is that the server will be able to authenticate the HTTP requests as coming from the client with whom it shares key K, and the server will be assured that the client's HTTP request could not be modified by a man-in-the-middle attacker.

C-7.8 Suppose a web client and web server for a popular shopping web site have performed a key exchange so that they are now sharing a secret session key. Describe a secure method for the web client to then navigate around various pages of the shopping site, optionally placing things into a shopping cart. Your solution is allowed to use one-way hash functions and pseudo-random number generators, but it cannot use HTTPS, so it does not need to achieve confidentiality. In any case, your solution should be resistant to HTTP session hijacking even from someone who can sniff all the packets.

Client can seed the random number generator with the secret key and include the next pseudo-random number in the sequence in each HTTP request, as well as a unique user id. The server can verify if this is the specified user by generating the random number with the shared secret key as seed and check if the numbers in sequence equal.