Paxson Spring 2011 — CS 161 Computer Security — Homework 4

Due: Monday May 2, at 11:59pm

Instructions. Submit your solution electronically *via your class account* by Monday May 2, at 11:59pm. You should upload a single file, `HW4.pdf`. Your writeup should include your name, your class account name (e.g., `cs161-xy`), your TA's name, the discussion section time where you want to pick up your graded homework, and "HW4" prominently on the first page. Use a legible font and clearly label each solution with the problem/subproblem to which it belongs. You *must* submit a PDF file; we will not accept other formats.

You must work on your own on this homework.

Updated 24Apr11: due date shifted three days later to Monday (with the consideration mentioned in the next paragraph). Also, some typos fixed.

Note: while this assignment is due on Monday May 2, you need to turn it in by the original due date of ***Friday April 29*** *if you want to assure that it will be graded and available for you to pick up several days before the final exam on May 12.*

Note: some of these problems look back to topics addressed earlier in class. Keep in mind that the final exam will be comprehensive across all topics.

Problem 1 ***DNSSEC*** **(20 points)**

DNSSEC (DNS Security Extensions) is designed to prevent network attacks such as DNS record spoofing and cache poisoning. When queried about a record that it possesses, such as when the DNSSEC server for `example.com` is queried about the IP address of `www.example.com`, the DNSSEC server returns with its answer an associated *signature*.

For the following, suppose that a user R (a resolver, in DNS parlance) sends a query Q to a DNSSEC server S, but all of the network traffic between R and S is visible to a network attacker N. The attacker N may send packets to R that appear to originate from S.

(a) Suppose that when queried for names that do not exist, DNSSEC servers such as S simply return "No Such Domain," the same as today's non-DNSSEC servers do. This reply has no associated signature.

Describe a possible attack that N can launch given this situation.

(b) Suppose now that when queried for a name Q that does not exist, S returns a signed statement "Q does not exist."

1. Describe a DoS attack that N can launch given this situation.

2. Describe a circumstance under which N can still launch the attack you sketched in the first part above; or explain why this attack no longer works.

(c) One approach for addressing the above considerations is to use NSEC Records. As mentioned in lecture, when using NSEC S can return a signed statement to the effect of "when sorted alphabetically, between the labels L_1 and L_2 there are no other labels." Then if the label L_3 in the query Q lexicographically falls between L_1 and L_2, this statement serves to inform R that indeed there's no information associated with L_3.

We discussed in lecture how NSEC has a shortcoming, which is that an attacker can use it to *enumerate* all of the labels in the given domain that do indeed exist. To counter this thread, in the April 6 section materials we introduced the NSEC3 Record, which is designed to prevent DNS responses from revealing unnecessary information. NSEC3 uses the lexicographic order of *hashed* labels, instead of their unhashed order. In response to a query without a matching record, NSEC3 returns the hashed names that come just before and just after the hash of the query.

Suppose that the server S has records for `a.example.com`, `b.example.com`, and `c.example.com`, but *not* for `abc.example.com`. In addition, assume that `a` hashes to 30, `b` to 10, `c` to 20, and `abc` to 15.

If the query Q from R is for `abc.example.com`, what will S return in response? Describe how R uses this to validate that `abc.example.com` indeed does not exist.

(d) In more detail, the way the hashes work in NSEC3 is they are computed as a function of the original name plus a *salt* and an *iteration parameter*, as follows:

```
Define H(x) to be the hash of x using the Hash Algorithm selected by
the NSEC3 RR, k to be the number of Iterations, and || to indicate
concatenation.  Then define:

   IH(salt, x, 0) = H(x || salt), and

   IH(salt, x, k) = H(IH(salt, x, k-1) || salt), if k > 0

Then the calculated hash of a name is

   IH(salt, name, iterations)
```

In an NSEC3 reply, the name of the hash function, the salt and the number of iterations are also included (that is, they are *visible* and assumed to be easily known). All replies from a given server use the same salt value and the same

number of iterations.

Suppose an attacker has a list of "names of interest," i.e., labels for which they want to know whether the given label is in a particular domain. If the attacker can get all of the NSEC3 responses for the particular domain, can they determine whether these names exist? If so, sketch how. If not, describe why not.

(e) What is the purpose of the *salt* in NSEC3 replies?

(f) What is the purpose of the *iteration parameter* in NSEC3 replies?

(g) The specification of NSEC3 also sets an upper bound on the *iteration parameter*. What threat does that protect against?

Problem 2 ***Covert Channels*** **(20 points)**

Consider a highly secured operating system that runs jobs for multiple users, but tries to assure strict isolation between the jobs, i.e., the jobs have no means to communicate with one another. The OS not only prohibits use of shared memory and any other forms of interprocess communication (pipes, sockets, shared memory, signals, use of the `ps` command), but also eliminates shared resources such as a global file descriptor pool.

The OS does, however, provide read-only access to a common file system. This file system does *not* make available "access time" information.

(a) Sketch how two cooperating processes could use the common file system to create a covert channel for communication. You can assume that the OS tries to optimize performance when accessing files by caching recently accessed blocks.

(b) In rough terms, what is the capacity of your covert channel? State any assumptions you make in your estimate.

(c) Suppose the OS eliminates the common read-only system in its entirety, though retains the performance technique of caching recently accessed blocks. Can two cooperating processes still communicate? If so, sketch how, and again estimate in rough terms the capacity of the channel. If not, then explain why communication is no longer possible.

Problem 3 ***Detecting Worms*** **(20 points)**

Assume that you are working for a security company that has to monitor a network link for worm traffic. The link connects a large site with the rest of the Internet, and always has lots of traffic on it.

Your company sells a monitoring box that can scan individual packets for fixed strings at very high speeds.

(a) Suppose that after careful analysis you discover that a particular TCP-based worm that you need to protect against generates traffic that always contains a fixed 4-byte sequence. You program your company's specialized hardware to generate an alarm whenever it detects a packet containing this 4-byte pattern.

Explain how this detector can exhibit false negatives.

(b) Propose an alternative architecture for your company's monitoring box that fixes the problem from part (a). Does your revised approach completely eliminate false negatives? Explain why or why not.

(c) Suppose benign network traffic is uniformly distributed in terms of its content. That is, every payload byte is chosen uniformly at random. Calculate the expected time until a 100 MB/sec link will cause a false alarm.

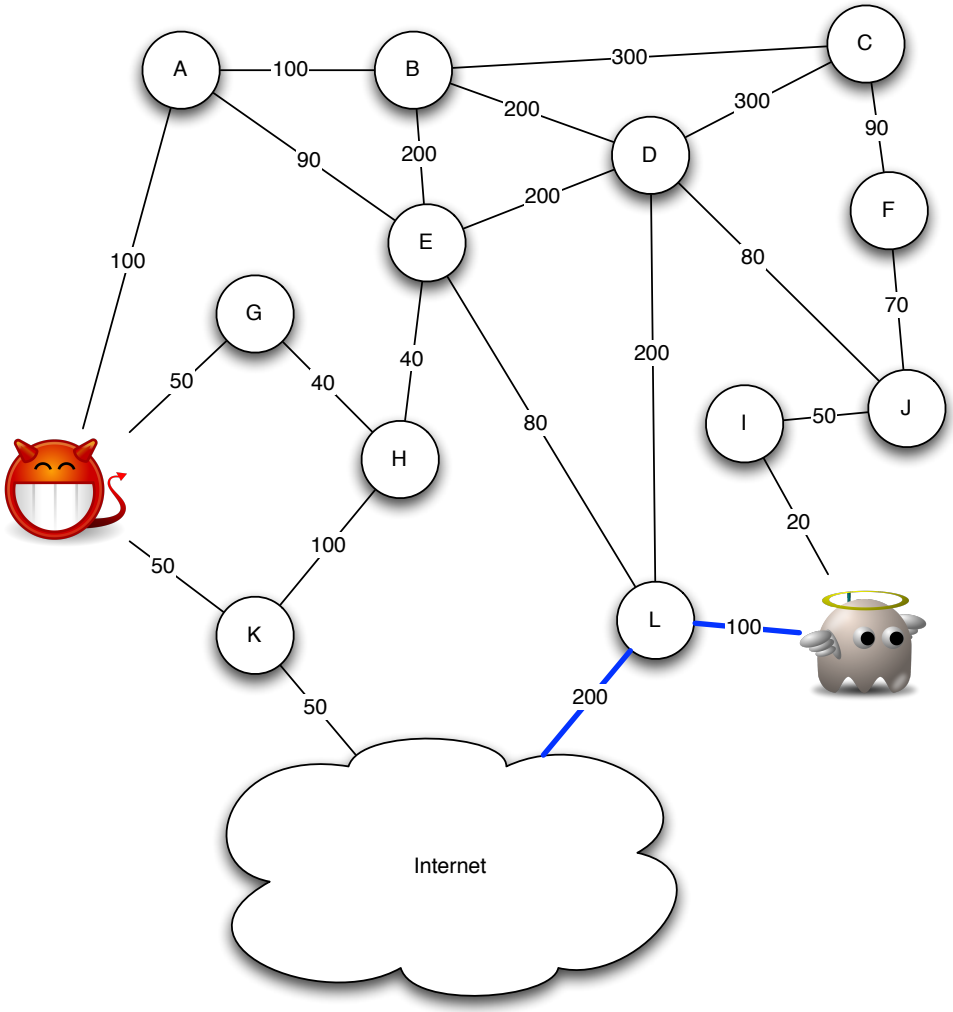
(d) Assume now that the signature length is 8 bytes. Calculate the expected time until a 10 GB/sec link will cause a false alarm.

(e) Explain a technique that worm authors can employ to try to ensure that their worms do not have many fixed-byte sequences.

Problem 4 ***Denial-of-Service Attacks*** **(20 points)**

The figure below shows the network of the country of Caltopia. It is quite a small country and contains only a dozen routers and two upstream connections to the rest of the global Internet. Each alphabetically labeled node corresponds to a router. The links between the nodes represent the maximum capacity in megabits/sec (Mbps). For example, the maximum bandwidth between G and H is 40 Mbps. Assume that if an attacker can fill such a link with traffic at that level in either direction, then the link becomes unusable for regular (benign) traffic.

The two players of interest are the attacker on the left and the victim on the right.



To route traffic between two nodes, the nodes always choose the shortest path according to the number of hops.[1] The routing is thus independent of the link capacity. An example routing from B to J would be the path B-D-J. In addition, assume that routing does not transit the "Internet" cloud in the diagram (that is, it won't go into the cloud and back out).

(a) The Internet connection of the victim is highlighted with bold links (also, blue, if viewed in color). Is there any means by which the attacker at the other end of the network can deny service to the victim in terms of cutting the victim off from the

[1] If there is a "tie" between two paths of the same length, a router resolves it by picking the path for which the next hop comes first alphabetically. For example, to get to node H the attacker will use the path through G rather than K.

global Internet by saturating the link from the victim to router L?

(b) Suppose that the victim's web server infrastructure can process 200,000 TCP connection requests per second. Assume that a single connection request (TCP initial SYN) requires 40 bytes to transmit. Can the attacker launch a sufficiently large SYN flood to overwhelm the victim's server? Explain why or why not.

(c) Suppose the victim configures routers I and L to block all incoming TCP packets with only the SYN bit set, but to allow inbound TCP packets with both SYN and ACK set. This denies all inbound connection requests to the victim. Can the attacker launch a SYN flood attack against the victim? If not, is the victim now safe from any type of TCP DoS?

(d) The *Slashdot effect* (also termed "flash crowds") is an unintended DoS attack by a massive number of visitors that click on a popular link in a news article, which causes them to load pages from a web server whose infrastructure lacks sufficient capacity to accommodate the large number of visitors. This popularity overloads the small site by either clogging its available bandwidth, or consuming all of the web server's resources.

After you graduate, your incredibly k00l startup becomes *slashdotted*. The Slashdot article contains a link to your company's lightly-provisioned website. How could you leverage this incoming traffic to launch a DoS attack against someone else's website?

Problem 5 ***Sandboxing*** **(20 points)**

Address Sandboxing aims to confine all memory reads and writes made by untrusted code to a pre-determined region ("segment") of memory. It is possible to enforce address sandboxing using an *inline reference monitor* technique called software fault isolation (SFI). SFI embeds an inline reference monitor to explicitly confine all memory accesses to refer to one segment in the process address space.

Consider a target architecture that has only two memory access instructions, load and store. `load [r1],r2` reads the value from memory pointed to by register `r1` into register `r2`, and `store [r1],r2` writes the value of register `r2` to memory pointed to by register `r1`. The target architecture has only one unconditional control transfer instruction, `jmp [r1]`, which transfers control to the address stored in r1. All instructions in the architecture are 8 bytes in length. You may assume other instructions exist if you need them to solve this question, but be sure to define them before using them in your answers.

(a) Suppose the SFI implementation employs the following checks for the enforcing address sandboxing policy: before each `load [r1],r2` and `store [r1],r2` instruction, *insert*:

```
r1 = r1 & 0x00000fff
r1 = r1 | 0xaaaaa000
```

The following checks are inserted before `jmp [r1]` instructions:

```
r1 = r1 & 0x00000fff
r1 = r1 | 0xddddd000
```

This implementation of the address sandboxing scheme aims to confine memory reads and writes to the memory range `0xaaaaa000`–`0xaaaaafff`. and the untrusted code with inline checks can only execute in the memory segment `0xddddd000`–`0xdddddfff`.

Explain why this implementation still allows malicious code to write to memory outside the intended range.

(b) Outline an alternate implementation that is secure, and explain why your proposed implementation is safe.