Name:_____ Student ID:_____

CSE 127 Computer Security
Spring Quarter, 2009
Final Exam

Instructor: Stefan Savage

Name _____
Student ID _____

Attention: This exam has 6 sections and you must answer all of them. Also, please write your name and student ID on each section. You have 170 minutes to complete the questions. As with any exam, you should read through the questions first and start with those that you are most comfortable with. Be sure to answer **all** parts of each question you answer. f you believe that you cannot answer a question without making some assumptions, state those assumptions in your answer. For partial credit, be sure to show how you arrived at your answer as well as the answer itself.

1	/8
2	/8
3	/8
4	/6
5	/8
6	/4
Total	/36

Name:_____ Student ID:_____

1. Quickies (8pts)

Short Answers (1 pt each)

a) Name one security benefit of using Network Address Translation (NAT)?

It prevents any hosts inside the NAT from being addressed directly. Similarly, it makes it easy to only allow outbound connections.

b) Why is spam e-mail typically delivered via botnets?

Because hosts become blacklisted once they send spam and thus spammers always need to have a fresh source of IPs to send with.

c) What does a stack cookie do?

It detects if an overflow of a stack buffer has overwritten stack meta-data [be generous with this one]

d) What is the difference between a polymorphic virus and a metamorphic virus?

A polymorphic virus has an encrypted body that is decrypted during execution. The actual plaintext virus code that eventually executes is static (the decryptor itself may be static or metamorphic). By contrast, a metamorphic virus, changes the particular instructions used by the virus each time it spreads.

True or False (1pt each)

T F Alice has a message to send to Bob and she wants to encrypt the message using public-key cryptography so that no one other than Bob can read it, she does so by using Bob's public key.

T **F** Properly used, a MAC can provide both confidentiality and authenticity.

T **F** One advantage of public-key cryptography is that, when properly implemented, it is much faster than symmetric key cryptography.

T F The SSL protocol is used to both authenticate Web servers and provide confidentiality for client transactions with them.

Name:_____ Student ID:_____

2. Passwords (8pts)

On June 7th, LinkedIn confirmed that it had experienced a data breach that likely compromised the e-mail addresses and passwords of 6.5 million of its users. This confirmation followed the posting of the password hashes for these users in a public forum. One criticism of LinkedIn is that they used *unsalted* password hashes. In this question we'll explore this criticism.

Assume that each stolen password record had two fields in it: [`user_email,SHA1(password)`] and that a user login would be verified by looking up the appropriate record based on user_email, and then checking if the corresponding hashed password field matched the SHA1 hash of the password inputted by the user trying to log in. By contrast, if LinkedIn had used a salted scheme, then each record would have had three fields: [`user_email,salt,SHA1(password+salt)`] and login verification would similarly require looking up the salt and using it when matching hashes. Given this:

a) Suppose the attacker's goal is to break *your* password via a dictionary attack. Does the lack of salting in LinkedIn's scheme make this goal substantially easier?

No. Because even were there a salt, the attacker knows what salt is used for a given user. Thus, the time to create a dictionary of hashes is pretty much the same in both schemes.

b) Suppose the attacker's goal is to break at least half of the passwords via a dictionary attack. Does the lack of salting in this scheme make this goal substantially easier?

Yes. Without salting one dictionary of hashes is sufficient for searching the entire set of users. With salting it will require a dictionary for each salt value seen.

c) Suppose you are contacted by the attacker and given a set of password hashes (that's it, no user_name, no salt). Assuming the hash function is known, is there a measurement you could make on order to infer if the hashes are likely salted or not?

Yes. Recall that some passwords are much more popular than others. For example, the password 123456 is used by at least 0.1% of all accounts. Thus, if you hash such passwords and they appear disproportionately in the list then you might infer that the list is not hashed. Similarly, even without doing a hash, if you sort the hashes by frequency, in an unsalted list you will expect that there is some hash that occurs with frequency ~ 0.1%, whereas in a salted list it will be ~0.1%/2^n where n is the size of the salt in bits.

d) It turns out that that 20% of LinkedIn users with Yahoo Mail e-mail addresses used the same password at LinkedIn as Yahoo. You learn that, unlinked LinkedIn, Yahoo salts its passwords. Should Yahoo be concerned about the LinkedIn breach or not?

Yes. For 20% of the Yahoo users in the LinkedIn breach, their user name and password is known to the attacker. Yahoo's salting helps mitigates a breach of their password database, but doesn't help at all in this case.

Name:_____ Student ID:_____

3. Costs (8pts)

Right after graduating from UCSD, on the basis of your experience in CSE 127 you are recruited to work for Google's little-known covert spy division. Your first assignment is to acquire Apple's secret plans for the iPhone 7 by any means necessary. Using the lock-picking skills you acquired in class, as well as your natural ninja-like stealth, you break into Apple's corporate headquarters and acquire the iPhone7.doc file. Unfortunately, you learn that the 100Kilobyte file is encrypted using the AES block cipher with 128-bit key. Faced with this problem, you have three strategies available for recovering the plaintext of the document

i) **Bribery**. Bribe an Apple employee who has access to the key used to encrypt the document. The employee is willing to provide you the key – for the right price. However, they are naturally nervous about being caught and hence they will only leak you 4 new bits of the key each week via a covert channel (at a cost of $100,000 per 4 bits).

ii) **Side-channel**. Call up the Q division at Google and get them working on a sophisticated side channel attack that will infer the bits of the document by remotely monitoring the power drawn by monitors inside Apple's Cupertino headquarters. Because of this attack's complexity it will provide only 1Kilobyte of the document each week (each week provides a new 1KB) and will incur an up-front cost of $1M to pay for the quantum lasers and mult-fractal machine learning technology.

iii) **Brute force**. Buy bots from a major Russian botnet operator at a cost of $1 per 2^{10} bots. Each bot can brute force 2^{40} keys per week (about 1.8 million keys per second). The supplier can sell you up to 2^{24} bots (approximately 17 million) until their supply is depleted.

You can use any of these strategies, alone or in combination.

a) Suppose you need to read the document in the quickest possible time. Which approach will do so, how long do you expect it to take and how much will it cost in the end?

Quickest approach: Bribe the employee for 16 weeks, learning 64 bits of the key. Then, buy 2^{24} bots and use them to brute-force the remaining 64 bits of key. This will take 17 weeks and cost $1,616,384. ($1,600,000 for the bribes and $16,384 for the bots). By contrast brute force would take 2^{64} weeks (the sun will explode first) and bribery alone will take 32 weeks (next best approach).

b) Suppose you only need to guarantee that you can read the document within 5 years (260 weeks) and thus you want to minimize the cost to do so. Which approach will do so, how long do you expect it to take and how much will it cost in the end?

Cheapest approach: Use the side channel. It will cost $1,000,000 and take 100 weeks. If we ignore the side channel, next cheapest method is to bribe the employee for 15 weeks, then brute-force for 16 weeks (but this costs 1.5M on bribes alone).

.

Name:_____ Student ID:_____

4. Code security (8 pts)

In the final class, I briefly mentioned that our group at UCSD has done a bunch of work looking at automotive security, including finding and remotely exploiting vulnerabilities in real cars. One of these vulnerabilities was with the CD player in a late model vehicle. We discovered a vulnerability whereby a particularly formatted CD could take control of the player and ultimately the entire vehicle. Based on this experience, consider the following C code:

```
/* Information about the current CD. */
struct cd {
   int numtracks;      /* The number of tracks on this disc. */
   int tracklen[16];   /* The length of each track on the disc, in seconds. */
   void (*notify)(struct cd *);   /* Call this whenever the CD info changes. */
};

struct cd *curcd = makestructcd();

/* Update the length of track number 'track'. */
void update_cdinfo(int track, int newtracklen) {
   if (track > 16)
      return;
   curcd->tracklen[track] = newtracklen;
   (curcd->notify)(curcd);
}
```

(Don't worry about `makestructcd()` it just allocates and initializes a `struct cd`.) Assume the adversary can arrange for `update_cdinfo()` to be called with whatever values of `track` and `newtracklen` he likes (those values may have been read directly off the CD, for instance). Integers and pointers are 32bits long. Answer the following questions about this code, concisely:

a) What is the security vulnerability in this code?

Answer 1: Buffer overrun (or array out-of-bounds error): if track=16, then this writes one past the end of the curcd->tracklen array.

Answer 2: Buffer overrun (or integer overflow error): if track is negative, the array dereference curcd->tracklen[track] writes outside the bounds of the curcd->tracklen array (it writes before the start of the array).

Either answer is ok

b) How could an attacker exploit this vulnerability to trigger the execution of malicious code? Describe how the attacker should choose the values of `track` and `newtracklen`.

Answer 1: Set track=16 and make newtracklen the address of malicious code loaded somewhere in the address space of this program. This will overwrite curcd->notify with a pointer to malicious code.

Answer 2: Set track to some large negative number chosen so that curcd->tracklen[track] references, e.g., a return address stored on the stack somewhere, and make newtracklen the address of malicious code loaded somewhere in the address space of this program. (There are many possible variations on this answer.)

Name:_____ Student ID:_____

5. Network protocols (8 pts)

A consortium of printer vendors have come up with a great new protocol to help users automatically discover the set of printers on their local network. In this protocol, when the user wants to print something, the user's computer automatically broadcasts a *Printer Discovery* packet. A *Printer Discovery* packet is a UDP packet whose destination address is the broadcast address, and whose source and destination port is 56184. Because this is a broadcast packet, every host on the local network will receive it. Printers constantly listen for *Printer Discovery* packets. Any time that they receive one, they immediately respond with a *Printer Announcement* packet. A *Printer Announcement* packet is a UDP packet whose destination address is the broadcast address, and whose source and destination port is 56185; its payload identifies the name of the printer, the printer's IP address, and any special options supported by the printer (e.g., 2-sided printing, color printing, 3D printing, etc). The *Printer Announcement* packet is broadcast to the entire network, so that other hosts on the local network can also learn about this printer. Whenever a machine receives a Printer Announcement packet, it checks that the source address of the packet matches the printer's IP address found in the payload. In case of a mismatch, it ignores the packet. Otherwise, it accepts the packet and adds this printer to its list of known printers. To accommodate changes in address assignment, if the machine's list of known printers already contains a printer with the same name, the machine overwrites the previous entry in its list with the information found in the newly received packet.

Victor the Victim is about to connect his laptop to a local switched Ethernet network. His laptop will use this printer discovery protocol to look for a printer, and then Victor will connect to one of the printers found in this way and send it a sensitive corporate document to be printed. Meanwhile, Hermione the Hacker is attached to this same network. Hermione has the ability to inject packets onto this network and to receive all broadcast packets, but she cannot eavesdrop on other traffic. The printers are in locked rooms that Hermione does not have access to, and Hermione has not been able to hack or access any of the machines or printers attached to this network, so her only hope is to attack the printer discovery protocol.

a) Can Hermione arrange to learn the contents of Victor's document, without physically accessing any of the printers? If yes, describe the attack, if no explain why the attack isn't possible.

Yes. Hermione can observe Victor's *Printer Discovery* packet and the real printers' *Printer Announcement* packets, then (before Victor prints the document) broadcast *Printer Announcement* packets containing Hermione's IP address but the name of the other printers. When Victor prints his document, he will send it to Hermione, and Hermione can see the contents of the document. Hermione can then optionally forward the document on to the printer so Victor doesn't notice anything amiss.

b) Can Hermione modify what is printed on the printer? In other words, Hermione wants to replace Victor's chosen document with something else Hermione has chosen, hopefully without Victor noticing. It's not acceptable if Victor's original document gets printed in addition to Hermione's replacement, because then Victor might notice and get suspicious. Can Hermione mount such an attack without physically accessing any printers? If yes, describe the attack, if not explain why the attack isn't possible.

Yes. Do the same as in (a), except modify the document before forwarding it on to the printer.

Name:_____ Student ID:_____

6. (8 pts) Worms/Malware/Spam

Consider two random scanning worms (i.e., like Code Red or Slammer) called A and B, that each select IP addresses to infect at random (out of roughly 4 billion IP addresses on the Internet). Worm A targets a vulnerability present in 1,000,000 hosts and each worm instance targets 10 IP random addresses per second. Worm B targets a different vulnerability present in 10,000 hosts, but targets 1000 IP random addresses per second.

a) Assuming both worms start at the same time, after one minute which worm do you expect will have compromised more additional hosts and why?

In the beginning of Worm A's growth the probability of finding a new host to infect is roughly 1M/4B (or 1/4000) and each second worm A will try 10 times for an expected rate of 1/40. Thus after one minute we expect worm A will have compromised one other host. By contrast, Worm B's expected growth rate will be 1/400… and we expect it will not have encountered another vulnerable host in one minute.

b) Same question, but after one day?

After one day, the Code Red worm infected virtually all 350k hosts scanning at a rate of approximate 10/second like Worm A. We expect that Worm A will do at least as well and will easily exceed the maximum number of infectable hosts in Worm B of 10,000.

c) Web mail providers, such as Yahoo, Hotmail and Gmail, automatically filter spam using automated techniques using both content and sender reputation. However, they also integrate user feedback as well. Thus, there is typically a button users can click labeled, "This is spam", when reading a message in your Inbox, as well as "This is not spam" when reviewing messages in your Junk folder (where it files messages it believes are spam). Such feedback could be used to directly train the underlying spam classifiers (e.g., messages labeled "is not spam" are instances of known good messages, while messages labeled "is spam" are known bad messages). However, there are real problems with doing this in practice. Explain what you think these problems might be.

An adversary can create fake accounts, send themselves spam then login and mark each spam message "this is not spam" to poison the training set for the classifier. Alternatively, they could send themselves legitimate messages and market it as spam thus increasing the overall false positive rate.

d) You encounter a new piece of malware. When it starts, it spawns a thread that executes: sleep(120) (effectively waiting for two minutes) after which it decrypts the malware payload and starts executing it. What anti-virus defense is the sleep call designed to work around?

This is designed to bypass generic decryption. The hope is that the AV program will refuse to simulate the malware's execution for two minutes and thus will miss the opportunity to see the malware's plaintext for matching.