CS 161
Computer Security

Dawn Song
Spring 2012

Midterm

Your Full Name: _____

Your Berkeley Email: _____

This is a closed-book midterm. You may not consult any lecture or written notes, cheatsheets, textbooks, etc. Calculators and computers are not permitted. Please write your answers in the spaces provided in the test. We will not grade anything on the back of an exam page unless we are clearly told on the front of the page to look there.

You have 80 minutes. There are 6 questions, of varying credit (62 points total). The questions are of varying difficulty, so avoid spending too long on any one question.

Do not turn this page until your instructor tells you to do so.

Question	Points	Total
Problem 1		6
Problem 2		10
Problem 3		9
Problem 4		8
Problem 5		10
Problem 6		18
Total		61

1. (6 points) Control Hijacking

   Indicate whether the statement is always valid. Indicate true or false, and give a one sentence explanation.
   Answer: +1 point for correct true false statement. +1 point for correct explanation.

   (a) (2 points) A stack canary prevents control hijacking from occurring.

   - ○ True. Reason: _____
   - ○ False. Reason: _____

   Answer: False. Canaries protect against stack based attacks by detecting when the return address is modified. This prevent buffer overflows, but does not prevent other exploits like exception handlers, pointer overwriting/subterfuge, heap exploits, etc.

   (b) (2 points) Consider the following program:

```
typedef void (*type_fp)(void);

void happy_function() {
  // something
}

int a(char *s) {
  type_fp hf = (type_fp)(&happy_function);
  char buf[16];
  strncpy(buf, s, 18);
  (*hf)();
  return 0;
}
```

   Assume that you control the input to the function $a()$. You can make the program behave incorrectly and jump to any arbitrary address.

   - ○ True. Reason: _____
   - ○ False. Reason: _____

   Answer: False. In this scenario, a maximum of 2 bytes of the function pointer can be overwritten (size 16 buffer on stack,

with size 18 parameter to strncpy). In order to jump to any arbitrary address, 4 bytes of address space are required to be overwritten.

(c) (2 points) Return oriented programming (arc-injection) is a viable technique to use to defeat stack canaries.

- ○ True. Reason: _____
- ○ False. Reason: _____

Answer: False. The canary is stored on the stack at a lower memory address than the return address. Meaning, when the buffer is overflowed to change the return address for an arc injection attack, the canary will also be over-written, causing the data injection to fail. Arc injection in general, only defeats non-executable stack protection.

2. (10 points) Symbolic Execution

Consider the following program:

```
1  void caller (int a, int b) {
2    int *ptr;
3
4    if (b > 0) {
5      ptr = getbuf(a,b);
6      if (ptr != NULL)
7        ptr[0] = 0;
8    }
9  }
10
11 int *getbuf (int x, int y) {
12
13   /* initialize all elements to zero */
14   int buf[20] = { 0 };
15   int z;
16
17   if (x > y)
18     return NULL;
19   if (x < 0)
20     z = -x;
21   else
22     z = x;
23   if (z < 20)
24     buf[z] = y;
25   return buf;
26 }
```

(a) (2 points) Consider the assignment at line 7. Is this assignment memory safe? Explain your reasoning in one sentence.

- Yes. Reason: _____
- No. Reason: _____

Answer: No. The pointer points to a local buffer allocated on stack; it's gone when the function returns.

(b) (4 points) Suppose we employ whitebox fuzzing to check if the program is vulnerable to a buffer overflow, i.e., applying dynamic sym-

bolic execution for automatic test case generation. In each run, white-box fuzzing creates new test cases for the parameters a and b to function `caller`. By converting the program statements into SSA (Static Single Assignment) form, write down the path constraints P on the symbolic inputs necessary to reach line number 24. Express P in terms of symbolic variables a_0, b_0, x_1, y_1 and z_1.

Answer:

$$(b_0 > 0) \tag{1}$$
$$\wedge\, (x_1 == a_0) \wedge (y_1 == b_0) \tag{2}$$
$$\wedge\, \neg(x_1 > y_1) \tag{3}$$
$$\wedge\, (z_1 == ((x_1 < 0)? - x_1 : x_1)) \tag{4}$$
$$\wedge\, (z_1 < 20) \tag{5}$$

1 pt for each non-equivalent clause: (1), (3), (4) and (5).

(c) (1 point) Write down the security assertion Q that you would insert at line 24 before the assignment to prevent a buffer overflow. Express Q in terms of symbolic variable z_1.

Answer: $z_1 >= 0$ **is sufficient.**

(d) (3 points) By solving the constraints $(P \wedge \neg Q)$ over 32-bit modulo arithmetic (`int` is represented in 32-bit two's complement), give an

instance of values for inputs a_0 and b_0 that are sufficient to cause a buffer overflow at line 24. You must provide concrete values free of macros or variables.

Answer: $a_0 == -2^{31}$

b_0 **can be any integer between 1 to** 2^{31}**-1 inclusive**

3. (9 points) Isolation, Least Privilege & Mobile Security

(a) (2 points) Dolphin Communications has come up with a brilliant idea for ensuring air gapping, named the SeaGap. SeaGap consists of a memory unit and several electronic switches. These switches are configured such that the memory could be connected only to the Internet or to your LAN, but never to both at the same time. When data arrives at one network port, the device would load up with application data, then flip 'safely' to the other network to disgorge its payload.

Does this achieve air Gap isolation?

- ○ Yes.
- ○ No.

Answer: No.

The two are still connected via the same computer on which SeaGap runs.

(b) (4 points) For each of the following security mechanisms, state whether they are access control or capability systems or both. If you answer 'both' write a short explanation for the same.

1. Safe that requires a user key: _____ **Answer: Capability**

2. Safe that checks your fingerprint and sees if you are in the list of allowed users: _____ **Answer: ACL**

3. Google Docs Sharing "Anyone with link" feature: _____ **Answer: Capability**

4. Google Docs sharing "Following users only": _____ **Answer: ACL**

(c) (1 point) The `ping` command, at a high level, consists of three modules: a module that sends the ICMP packets, another module to receive the responses and a third module to show the output to the user.

`ping` runs as a monolithic process. Since sending the packets above requires root privileges, `ping` is a setuid binary that elevates to root as soon as it starts, and then sends the ICMP packets, receives the response and shows the output.

What security principle does this design violate?

- ◦ Confinement Principle
- ◦ Complete Mediation
- ◦ Principle of Least Privilege
- ◦ Low Coupling Design

Answer: Principle of Least Privilege.

Only sending the packet requires root privilege: the code that parses the response and shows the output doesn't need to run as root.

(d) (2 points) Which one of the following mechanisms is **NOT** an integral component required for Android application isolation?(**circle one**)

- ◦ Application code signing
- ◦ Android permission system
- ◦ Linux users
- ◦ Linux process isolation

Answer: Android permission system

4. (8 points) Trusted Computing

   (a) (1 point) How is a TPM implemented? (**circle one**)

   - Entirely in software.
   - In the BIOS firmware.
   - As a hardware component in the system.
   - Using the SKINIT x86 instruction.
   - As a cloud service.

   Answer: As a hardware component in the system.

   (b) (1 point) Suppose that BIOS code is updated by a firmware update. How would the system enable access to blobs previously sealed to the current BIOS version? (**circle one**)

   - It is not possible to patch the BIOS in this architecture.
   - The patch process must re-seal all blobs with new PCR values.
   - All blobs must be decrypted and stored in cleartext.
   - The TPM will decrypt old blobs even after the update.

   Answer: The patch process must re-seal all blobs with new PCR values.

   (c) (1 point) In BitLocker, what is the purpose of the boot-time PIN or USB key? (**circle one**)

   - To annoy the user.
   - So that if the machine is stolen, the attacker cannot decrypt the disk.
   - So that malware cannot change the OS loader.
   - To prevent moving the disk to another machine.

   Answer: So that if the machine is stolen, the attacker cannot decrypt the disk.

   (d) (1 point) A TPM can be used to speed up hard drive encryption (e.g., in BitLocker).

   - True
   - False

   Answer: False. It is used to provide TRUST (i.e., storing secrets and attesting the values of PCR's. TPM's are in fact

extremely slow and if they are used for hard drive encryption (assuming this is even possible), they will actually slow it down a lot.

(e) (1 point) Trusted boot can be used to detect that the operating system kernel has been modified by a virus.

- ○ True
- ○ False

Answer: This question was not graded so everyone will receive 1 point for this question regardless of their answer.

(f) (1 point) Trusted boot can be used to detect that the hardware has been tampered with.

- ○ True
- ○ False

Answer: False

(g) (1 point) A user with root/administrator privileges can read the internal memory of a TPM, but a user without root/administrator privileges cannot.

- ○ True
- ○ False

Answer: False

(h) (1 point) With DRTM (Dynamic Root of Trust Measurement), the BIOS is measured (i.e., its cryptographic hash is computed and stored in a PCR).

- ○ True
- ○ False

Answer: False

5. (10 points) Cryptography

(a) (1 point) There is a mathematical proof that factoring large numbers is computationally infeasible (i.e., it takes too much time).

- ○ True
- ○ False

Answer: False. Factoring is considered hard because it is a well known problem that no one knows how to solve efficiently.

(b) (1 point) How does a MAC (Message Authentication Code) differ from symmetric encryption? Note: To invert means to compute the input given the output. (**circle all that apply**)

- ○ It doesn't.
- ○ A MAC has a pair of keys (public and private) and symmetric encryption does not.
- ○ A MAC has a single key and symmetric encryption does not.
- ○ Symmetric encryption has a pair of keys (public and private) a MAC does not.
- ○ Symmetric encryption has a single key and a MAC does not.
- ○ Symmetric encryption has a fixed-size output and a MAC does not.
- ○ A MAC has a fixed-size output and symmetric encryption does not.
- ○ Given the key(s), it is always feasible to invert a MAC, but it is not always feasible to invert symmetric encryption.
- ○ Given the key(s), it is always feasible to invert symmetric encryption, but it is not always feasible to invert a MAC.

Answer: "Symmetric encryption has a fixed-size output and a MAC does not." and "Given the key(s), it is always feasible to invert symmetric encryption, but it is not always feasible to invert a MAC."

(c) (2 points) How does a MAC (Message Authentication Code) differ from a cryptographic hash function? Note: To invert means to compute the input given the output. (**circle all that apply**)

- ○ It doesn't.

- ◦ A MAC has a pair of keys (public and private) and a cryptographic hash function does not.
- ◦ A MAC has a single key and a cryptographic hash function does not.
- ◦ A cryptographic hash function has a pair of keys (public and private) a MAC does not.
- ◦ A cryptographic hash function has a single key and a MAC does not.
- ◦ A cryptographic hash function has a fixed-size output and a MAC does not.
- ◦ A MAC has a fixed-size output and a cryptographic hash function does not.
- ◦ Given the key(s), it is always feasible to invert a MAC, but it is always not feasible to invert a cryptographic hash function.
- ◦ Given the key(s), it is always feasible to invert a cryptographic hash function, but it is not always feasible to invert a MAC.

Answer: A MAC has a single key and a cryptographic hash function does not.

(d) (1 point) If I encrypt a message, then I don't need to authenticate it to prevent against active attacks.

- ◦ True
- ◦ False

Answer: False. If the message is not authenticated, the attacker can just replace it with another message or a random garbage message.

(e) (2 points) Suppose Alice wants to send a message to Bob so that only Bob can read it and so that Bob gets the correct message. They engage in the following protocol:

1. Bob sends Alice his public key.
2. Alice uses an asymmetric encryption algorithm to encrypt the message with the public key she just received.
3. Alice sends the ciphertext to Bob.
4. Bob decrypts the ciphertext he receives with his private key to obtain a message.

Which of the following are true? (**circle all that apply**)

- This protocol is secure against passive adversaries (eavesdroppers).
- This protocol is secure against active adversaries (man in the middle).
- None of the above.

Answer: This protocol is secure against passive adversaries (eavesdroppers) only. An active adversary can send his public key to Bob instead of Alice's public key. Then Bob will send out a message encrypted with the adversary's public key and the adversary can use his own private key to decrypt it and read the message. The adversary can then encrypt the message (or some other maliciously chosen message) using Alice's public key and send it to Alice. Look at certificate authority topic in the cryptography notes for a similar attack.

(f) (3 points) Suppose that you have a game installed on your laptop. The game periodically downloads executable updates from `http://<game-website>/updates/`. You now bring your laptop to class and connect it to the AirBears WiFi network. Note that AirBears is susceptible to man-in-the middle attacks. In order to prevent your computer from being compromised, when the game downloads an update `http://<game-website>/updates/updateX.exe` which of the following can the game do? (**circle all that apply**)

- Verify it against the digital signature stored in `http://<game-website>/updates/updateX-signature.txt` using the game company's public key that is already embedded in the game's code.
- Compute a MAC of updateX.exe and verify that it matches the MAC stored in `http://<game-website>/updates/updateX-mac.txt` using a MAC key that is already embedded in the game's code.
- Compute a cryptographic hash of updateX.exe and verify that it matches the cryptographic hash stored in `http://<game-website>/updates/updateX-hash.txt`
- It is not necessary to perform any cryptographic operations because it is not possible to perform man-in-the-middle attacks against HTTP.

Answer: Verify it against the digital signature stored in

`http://<game-website>/updates/updateX-signature.txt` using the game company's public key that is already embedded in the game's code. A MAC won't work because it is a symmetric algorithm and the game would have to embed the MAC key into the software for verification, and an attacker can then extract the key out of the software and use it generate a MAC of a malicious update. A cryptographic hash doesn't use keys so an attacker can generate the cryptographic hash key of a malicious update. HTTP is indeed susceptible to man-in-the-middle attacks.

6. (18 points) Web Security

   (a) (2 points) When visiting a website, such as a banks website, which of the following is a necessary part of preventing a man-in-the-middle attack?

   (a) An HTTPS connection
   (b) A security image
   (c) A CAPTCHA

   - (a) only
   - (b) only
   - (c) only
   - Both (a) and (b)
   - Both (b) and (c)

   Answer: (a) only

   (b) (1 point) In the following PHP code, in which line is there a potential XSS attack, assuming all sanitizer functions work correctly and all variables are user inputs?

```
1  <?php
2  echo '<p>Hello, ' . sanitizeHTML($username) . '</p>';
3  echo '<p>The homepage for user id ' .
4       sanitizeNumber($userid) . ' is:</p>';
5  echo '<p><a href= ' . sanitizeHTML($homepage).
6       '  >homepage</a></p>';
7  echo '<p><a href= myprofile.php  >' .
8       'Return to profile of ' .
9       sanitizeHTML($username) .
10      '.</a></p>';
11 ?php>
```

   - Line 2
   - Line 4
   - Line 5
   - Line 9
   - There is no XSS

   Answer: Line 5

(c) (4 points) In the `trusted.com` website, there are a number of references to external URLs at `untrusted.com`. For each of the following HTML elements that appear in the `trusted.com` website, *when the external resource is downloaded*, specify whether it is executed in the `trusted.com` or the `untrusted.com` origin.

1. `<a href="untrusted.com">` _____ **Answer: untrusted.com**
2. `<script src="untrusted.com">` _____ **Answer: trusted.com**
3. `<iframe src="untrusted.com">` _____ **Answer: untrusted.com**
4. `<style src="untrusted.com">` _____ **Answer: trusted.com**

(d) (2 points) You are visiting a banking website, `http://www.americasbank.com`. After logging in, a session is established with the server with a random 8 bit session ID in the cookie. Unfortunately, Mallory, a network attacker, is able to hijack your session with the bank and transfer out a large sum of money. Which of the following changes does the bank need to do to prevent such attacks in the future and provide the most flexibility? **Circle all that apply (leave blank for none). Point awarded only if all the correct options (and no others) are circled**.

- Use SSL/TLS.
- Increase the random session ID length.
- Check the IP address of the connection. If it is different from the previous IP address used with the given session ID, reject the connection.
- Create a new, random session ID every 5 minutes.
- Require the user to change their password at least once a month.

Answer: Increase random session ID length and Use SSL/TLS. +2 point for both, 0 points otherwise.

(e) (1 point) Are the following URIs same origin?

1. http://www.example.com:80/index.html
2. http://www.example.com/index.html

- Yes

- No

Answer: Yes

(f) (1 point) I go to a page on `http://www.example.net` and log in. The person who wrote the login page is my friend, and I know he always makes sure to set the form action (the target uri for form submission) to an HTTPS URI. Which one of the following options is correct? **Circle all that apply (leave blank for none). Point awarded only if all the correct options (and no others) are circled.**

- A network attacker cannot read my password since it is always sent to an HTTPS URI
- A network attacker can read my password because the form is submitted using a HTTP GET, which means the password is sent as part of the URI.
- A malware attacker can get my password.
- A web attacker can read my passwords using framebusting.

Answer: A malware attacker can get my password.

(g) (2 points) A web application firewall is a software program that sits on the network, next to the web application server and looks at all HTTP Requests going to the server. It is used to detect XSS attempts. Which of the following attack attempts could be detected by a web application firewall interposing on all requests to the `www.example.com` web server? **Circle all that apply (leave blank for none). Point awarded only if all the correct options (and no others) are circled**.

1. `http://www.example.com/postComment.php` with POST body `<script>doEvil()</script>`
2. `http://www.example.com/post.php?comment=<script>doEvil()</script>`
3. `http://www.example.com/search.php#!?=in=db&query=<script>doEvil</script>`
4. `http://blog.example.com/post.php?comment=<script>doEvil()</script>`

Answer: 1 and 2 only. Superfluous options gives zero points.

(h) (2 points) Prof. Evil provides all the members of CalTopia access to Zion, the centralized servers holding a distributed Badoop Filesystem for storing data. Users can ssh in and see their files and/or create new files. The linux kernel ensures the permissions setup. For example, files of `/home/profevil/` are all readable only by the `profevil` user and not by `minion420` user. Prof. Evil also creates a web based UI to view existing files. In this website, the UI requires that you login with your username and password. The web server runs as group `www`, and all user files are given group `www`. The server validates the login credentials with the OS. Finally, the web app looks up the owner of the file in question and makes sure the owner matches the logged in user. For example, if a file is not readable by `minion420`, then the WebUI will refuse to display it to him. Which of the following is correct regarding the check that the owner of the file matches the logged in user? **Circle all that apply (leave blank for none) Point awarded only if all the correct options (and no others) are circled**.

- The server code doesnt need to do this; the OS kernel takes care of it automatically via the permissions setup. They should remove the check for efficiency.
- The server code doesnt really need to do this, but its a good defense in depth mechanism.
- The server code needs to do this, as otherwise minion420 could read profevils files.
- The server code needs to do this because the OS kernels implementation might have a bug.

Answer: Only option 3 receives +2 points. Any other option, or selecting superfluous options results in zero points.

(i) (3 points) `BCS.com` wants to add social networking to its website using gracebook.com. For this, it needs to accept post-messages from gracebook.com subdomains. The following code does this check:

```
window.onmessage=function(e){
if(e.origin.indexOf('.gracebook.com') != -1){
//trust the message
}
```

The String `indexOf` method is defined as:

> The indexOf method returns the index within the calling String object of the first occurrence of the specified value, returns -1 if the value is not found.

We only want to accept messages from gracebook.com and all its subdomains. Is the check sufficient? If not, give a counter example (i.e., a possibly attacker controlled domain that will be trusted). **Answer: +1 point for no.**

+2 point for a correct example. One possible answer is `:foo.gracebook.com.attacker.com`