CS 155: Spring 2015
June 5, 2015

CS 155 Final Exam

This exam is open book and open notes. You may use course notes and documents that you have stored on a laptop, but you may NOT use the network connection on your laptop in any way, especially not to search the web or communicate with a friend. **You have 2 hours.**

Print your name legibly and sign and abide by the honor code written below. All of the intended answers may be written in the space provided. You may use the back of a page for scratch work. If you use the back side of a page to write part of your answer, be sure to mark your answer clearly.

The following is a statement of the Stanford University Honor Code:

A. *The Honor Code is an undertaking of the students, individually and collectively:*

   (1) *that they will not give or receive aid in examinations; that they will not give or receive unpermitted aid in class work, in the preparation of reports, or in any other work that is to be used by the instructor as the basis of grading;*

   (2) *that they will do their share and take an active part in seeing to it that others as well as themselves uphold the spirit and letter of the Honor Code.*

B. *The faculty on its part manifests its confidence in the honor of its students by refraining from proctoring examinations and from taking unusual and unreasonable precautions to prevent the forms of dishonesty mentioned above. The faculty will also avoid, as far as practicable, academic procedures that create temptations to violate the Honor Code.*

C. *While the faculty alone has the right and obligation to set academic requirements, the students and faculty will work together to establish optimal conditions for honorable academic work.*

I acknowledge and accept the Honor Code.

____
(Signature)

______________________ *(SUNet ID)*

______________________ *(Print your name,* legibly!*)*

□ **GRADUATING?**

Prob	# 1	# 2	# 3	# 4	# 5	# 6	# 7	Total
Score								
Max	11	20	11	12	10	10	12	86

1. (*11 points*) .... True or False

______ (a) Control hijacking attacks can give an attacker shell access to a vulnerable system.

______ (b) ASLR is designed to make it harder to mount XSS attacks.

______ (c) System call interposition is a method to speed up the processing of system calls in modern operating systems.

______ (d) The Same Origin Policy used for the DOM is the same as the Same Origin Policy applied to cookies.

______ (e) A CSRF vulnerability at a bank has the following consequence: a malicious web site can issue requests to the bank on behalf of users visiting the malicious web site.

______ (f) A collision resistant hash function is a hash function designed to prevent collisions in self-driving cars.

______ (g) The TLS session setup protocol uses only symmetric-key cryptography.

______ (h) Pinning the `gmail.com` certificate in the browser is done so that an attacker who obtains a certificate for `gmail.com` cannot masquerade as the Gmail server.

______ (i) Small sites can defend themselves against SYN floods by hiding behind a large 3rd party proxy service that only forwards completed TCP connections to the site.

______ (j) The principle of least privilege says that each system component needs to be given the privileges it needs to do its job, but nothing more.

______ (k) Suppose an attacker injects a spurious BGP route that affects traffic from a bank site in the US to a browser located elsewhere (but not in the reverse direction). The attacker can use the route to read a cleartext HTTP response from the bank to the browser even if the response is sent over HTTPS from the bank to that browser.

2. (*20 points*) .... Questions From All Over With a Short Answer

(a) (*4 points*) Briefly explain the purpose of ASLR.

(b) (*4 points*) Briefly explain why HSTS prevents `ssl_strip` attacks.

(c) (*4 points*) Suppose a web page `/site.com/index.php` contains the following PHP script: `<?php echo "Hello ".$_GET["name"]; ?>`
What vulnerability does this cause? Give an example exploit URL (do not URL encode your example).

(d) (*4 points*) On some laptops, a USB device plugged into the USB port can write to any physical memory location using Direct Memory Access (DMA). Explain why this is a security risk. What attacks are possible?

(e) (*4 points*) It is possible to download an employee's W-2 tax form using the employee's 9-digit social security number and an 8-digit PIN assigned by the employer. (This is really true.) Assume that each employer assigns PINs truly at random. One of the sites providing this service has a security page that says, "This site has security measures in place to protect against the loss, misuse or alteration of the information under our control. Our site make extensive use of Secure Sockets Layer (SSL) encryption ... In addition, we take great care to safeguard all information that is transmitted to us and stored by us on our computers, servers, and databases. This protection includes the use of network intrusion prevention and detection technology as well as other industry accepted security practices. We protect the physical location of our systems through the use of access control and monitoring technologies ..."

Question: How secure is your W-2 form? What is the simplest attack enabling an attacker to obtain your W-2 form, assuming he knows your social security number but not your PIN? How would you defend against the attack you propose?

3. (*11 points*) .... Control flow guard

(a) (*3 points*) Suppose a function pointer `fp` is allocated on the stack. Explain how a stack buffer overflow can let an attacker mount a return-oriented-programming (ROP) attack despite the presence of stack canaries and DEP. Make sure the attack you describe is an ROP attack. Throughout this question you may assume that ASLR is not used.

(b) (*4 points*) Windows 10 introduces a mechanism called control flow guard (CFG). CFG embeds a bitmap in the executable file that contains one bit for every eight bytes of executable code. The bit is set to 1 if the corresponding eight bytes of code contain the entry point to a function and is set to 0 otherwise. Before a function pointer `fp` is called, the program code checks that the address stored in `fp` corresponds to a bit in the CFG bitmap that is set to 1. If not, the program is terminated. Does this check prevent your attack from part (a)? If so explain why. If not explain why not.

(c) (*4 points*) Does CFG ensure that in the example from part (a) control flow always proceeds to where the program intended? That is, can an attacker cause the program to jump to the wrong location without being terminated? If so, please provide example pseudo-code where this could lead to an attack. If not, explain why not.

4. (*12 points*) .... Packet fragmentation and firewalls

The IP protocol supports fragmentation, which allows a packet to be broken into smaller fragments as needed and re-assembled when the fragments reach their destination. When a packet is fragmented it is assigned a 16-bit packet ID; each fragment is identified by its offset within the original packet. The fragments travel to the destination as separate packets. At the destination they are grouped by their packet ID and assembled into a complete packet using the packet offset of each fragment.

The IP packet format is often drawn with 32 bits per line:

<table>
<tr><td>Version</td><td>IP HL</td><td>TOS</td><td colspan="3">Total length</td></tr>
<tr><td colspan="3">Identification</td><td>Flags</td><td colspan="2">Fragment offset</td></tr>
<tr><td colspan="2">TTL</td><td>Protocol</td><td colspan="3">Header Checksum</td></tr>
<tr><td colspan="6">Source Address</td></tr>
<tr><td colspan="6">Destination Address</td></tr>
<tr><td colspan="5">Options</td><td>Pad</td></tr>
<tr><td colspan="6">Data</td></tr>
</table>

The Identification field is 16 bits and is a value assigned by the sender of an IP datagram to aid in reassembling the fragments of a datagram. The Protocol field specifies the next-layer protocol (e.g., TCP, UDP, ICMP) and is “byte 9” if we start the packet with “byte 0”. The fragment offset is 13 bits and indicates where a fragment belongs in the original IP datagram. This value is a multiple of eight. In the flags field of the IP header, bit 1 is the DF bit (0 = “may fragment,” 1 = “don’t fragment”). Bit 2 is the MF bit (0 = “last fragment,” 1 = “more fragments”).

If this is a TCP packet, for example, then the TCP header is contained in the portion of the packet shown here as “Data”. A specification saying which higher level protocol (e.g., HTTP, SMTP) is used is also in the “Data” portion of the IP format.

(a) (*4 points*) Suppose you want to install a stateless packet filter and block incoming SMTP packets but allow HTTP packets. Why is fragmenting a problem?

(b) (*4 points*) Suppose you are building a possibly stateful or application layer filtering engine. Your goals include: (i) block incoming SMTP traffic, but (ii) allow HTTP traffic. We noted in class that when fragments with overlapping segments are reassembled at the destination, the results can vary from OS to OS. One OS may use the data from the last received fragment for the overlapping segment while another OS may use the data in the first fragment received. Give an example of how overlapping fragments can cause the filtering engine to incorrectly allow some SMTP traffic through.

(c) (*4 points*) How should your filtering engine designed to meet the requirements stated in the part (b) handle overlapping fragments?

5. (*10 points*) .... The .bank Top Level DNS Domain

The .bank top-level-domain (TLD) is scheduled to open to the general public in 2015. This top level domain is only open to verified bank organizations, through a process supported by the American Bankers Association (ABA) and BITS, the technology policy arm of the Financial Services Roundtable. Under this new plan, Wells Fargo Bank could be assigned wellsfargo.bank and use this in place of wellsfargo.com.

Financial institutions that use the .bank TLD will be required to adhere to a list of requirements, including:

(i) Domain names will not be activated or resolve in the .bank DNS unless they match the trade name of a bank that registers them and are not deemed to be deceptive or confusing.

(ii) DNSSEC must be deployed at each zone and subsequent sub-zones for domains that resolve in the .bank DNS.

(iii) DNS Resource Records (e.g., CNAME, DNAME, SRV) are prohibited from aliasing to DNS records outside of the secure zone.

This question asks you to understand and explain some of the reasons for these .bank requirements.

(a) (*3 points*) What general protections does DNSSEC provide?

(b) (*3 points*) Why might the banking industry consider using DNSSEC with *opt-out* for .bank? (the reasons may not be for improved security)

(c) (*4 points*) Why do the regulations require DNSSEC at all sub-zones?

6. (*10 points*) .... Cross Site Script Inclusion (XSSI) Attacks

In this problem we look at a common Web vulnerability. Consider a banking web site `bank.com`. After login the user is taken to a user information page

```
https://bank.com/accountInfo.html
```

that shows the user's account balances. `accountInfo.html` is a static page: it contains the page layout, but no user data. Towards the bottom of the page a script is included as

```
<script src="//bank.com/userdata.js">                (1)
```

The contents of `userdata.js` is as follows:

```
displayData({"name": "John Doe",
             "AccountNumber":  12345,
             "Balance": 45})
```

The function `displayData` is defined in `accountInfo.html` and uses the provided data to populate the page with user data.

The script `userdata.js` is generated dynamically and is the only part of the page that contains user data. Everything else is static content. Keep in mind that line (1) causes the script `userdata.js` to be executed in the context of the page that includes it.

Suppose that after the user logs in to his or her account at `bank.com` the site stores the user's session token in a browser cookie.

(a) (*6 points*) Consider user `John Doe` who logs into his account at `bank.com` and then visits the URL `https://evil.com/`. Explain how the page at `evil.com` can cause all of John Doe's data to be sent to `evil.com`. Please provide the code contained in the page at `evil.com`

(b) (*4 points*) How would you keep `accountInfo.html` as a static page, but prevent the attack from part (a)? You need only change line (1) and `userdata.js`. Make sure to explain why your defense prevents the attack.
Hint: Try loading the user's data in a way that gives `bank.com` access to the data, but does not give `evil.com` access. In particular, `userdata.js` need not be a Javascript file.

7. (*12 points*) .... Email spoofing

Sender Policy Framework (SPF) is a simple email validation system designed to detect email spoofing. It works as follows: a site like `gmail.com` publishes in its DNS record an SPF entry specifying all the IP addresses that can send email on behalf of Gmail.

These SPF records are used as follows. An email server, say `mail.stanford.edu`, receives an email claiming to be from Gmail. Let P be the IP address of the server from which the incoming email is sent. The `mail.stanford.edu` server looks up the SPF record for `gmail.com` and rejects the incoming email if P is not on the list of authorized IP addresses who can send email on behalf of Gmail.

SPF is designed to prevent non-Gmail machines from sending email claiming to be from Gmail. Recall that the mail protocol SMTP runs over TCP: the sending server connects to the recipient's TCP port 25 and transmits the email. Establishing the TCP connection requires a 3-way handshake.

(a) (*2 points*) Without SPF, any machine on the Internet can send mail pretending to come from a Gmail user. Name one way this is widely used maliciously.

(b) (*3 points*) Suppose `mail.stanford.edu` always chooses its initial TCP sequence number as 0. Explain how this makes it possible for an attacker to completely bypass the SPF check. The attacker need only send packets to `mail.stanford.edu`. No eavesdropping is needed.

(c) (*3 points*) The attack from part (b) is not possible if `mail.stanford.edu` chooses a random initial TCP sequence number. Does this mean that SPF prevents every non-Gmail machine from sending mail on behalf of Gmail to `mail.stanford.edu`? (*Hint:* What if the attacker can eavesdrop on packets sent from `mail.stanford.edu`?)

(d) (*4 points*) Another spoofing defense called DKIM has Gmail digitally sign every outgoing email. Gmail publishes its public verification key in the `gmail.com` DNS record. When `mail.stanford.edu` receives an email claiming to be from Gmail it first fetches Gmail's public key from DNS, verifies the signature on the incoming email, and accepts the mail only if the signature verifies. Does DKIM prevent a network attacker from sending mail on behalf of Gmail users? If so explain why, if not explain why not.