Final exam in

Web Security EITF05

Department of Electrical and Information Technology
Lund University

October 19th, 2011, 14.00-19.00

- You may answer in either Swedish or English.
- If any data is lacking, make (and state) reasonable assumptions.
- Use legible hand writing. If your answers cannot be read, you will receive zero points on that problem.
- Grading is done as follows.
  Grade 3 = 20–29 points,
  Grade 4 = 30–39 points,
  Grade 5 = 40–50 points.

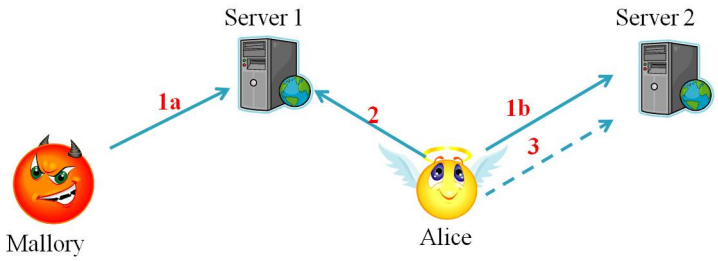
Good luck!

Martin, Paul & Christopher

Problem 1. You are an evil attacker that aims at bringing down the entire Internet. What should you reasonably direct your efforts towards, and which attack method should you select? Briefly explain how your selected attack method works and what you can do to maximize the success probability of your attack.

Answer
Attacking the root servers with a DNS amplification attack is one reasonable option. Using a bot net for a large scale attack will be necessary for this project to succeed, and using DNSSEC-enabled DNS servers will also increase the success probability. (3 points)

Problem 2. Explain how a CSRF attack works. Also, name at least one countermeasure. You may refer to the picture below.



Answer
1a: Mallory puts a script on webpage on Server 1.
1b: Alice logs in to Server 2 and obtains a session cookie.
The order of steps 1a and 1b is irrelevant.
2: Alice visits webpage on Server 1.
3: Mallory's script is executed on Alice's computer, sending a request to perform some action on Server 2.

CSRF protection on Server 2:

- Allowing only POST requests
- Requiring session ID to be sent in POST body or GET string as well as in the cookie.
- Check that referrer is as expected
- Make user reauthenticate before certain requests

CSRF protection on user-side:

- signing off

(3 points)

Problem 3. How can spammers bypass greylisting and nolisting?

Answer
A greylisting mailserver will look at (client ip, sender address, recipient address) and temporarily reject the message if the above triplet is previously unused. If recently used, it will accept the message. The spammer can implement a retry.
In nolisting, the first mail server listed in the MX record is non-existing. This can be circumvented by sending to the second mail server if the first one does not respond.
All-in-all, both defense mechanisms can be bypassed by following the protocol. (3 points)

Problem 4. Explain how a DNS cache poisoning attack can realize a man-in-the-middle attack.

Answer
In the attack, the adversary injects fake anwers to a query hoping that the querying server will accept the IP in the answer as belonging to the queried name. If it accepts, the server will cache this answer and all subsequent questions will be answered with the wrong IP. Assume that IP_a corresponds to www.a.se but an attacker can poison a DNS cache so that queries to www.a.se instead are answered with IP_b, a computer controlled by the attacker. Then all HTTP requests to www.a.se are sent to IP_b and these can be intercepted by the attacker, read/modified, and then forwarded to IP_a. This could be used e.g., to eavesdrop passwords sent in clear, perform an SSL man-in-the-middle attack if anonymous Diffie-Hellman is used etc. (3 points)

Problem 5. Give a regular expression that checks if an input is a number and that the number is divisible by 4.
Hint: You may assume that the number has at least two digits.

Answer
`^[0-9]*([02468][048]|[13579][26])$` (3 points)

Problem 6. Give the Base64 encoding of the word "DKIM".

Hint: Decimal representation of ASCII characters is given by:

$$\text{A} = 65, \text{B} = 66, \ldots, \text{Z} = 90, \text{a} = 97, \text{b} = 98, \ldots, \text{z} = 122$$

The Base64 alphabet is:

$$0 = \text{A}, \ldots, 25 = \text{Z}, 26 = \text{a}, \ldots, 51 = \text{z}, 52 = 0, 53 = 1, \ldots, 61 = 9, 62 = +, 63 = /$$

Answer
REtJTQ== (3 points)

Problem 7. Tor is a so called low-latency design for anonymous communication.

a) What is meant by this?

b) What drawbacks do low-latency designs have and how are they compensated?

Answer
a) A low latency designs means that messages entering a Mix (or a router), leaves soon afterwards with very short delay.
b) The drawback is that the anonymity set is necessarily quite small. Compensation include using several mixes, have a high traffic volume, use synthetic traffic, and taking advantage of the small anonymity set that still exists, mixing packets that arrive within a short time period. (3 points)

Problem 8. In a remote file inclusion attack, the adversary tricks the PHP script to read a remote file and interpret its content. However, according to the same-origin policy one origin cannot read resources in other origins. Why doesn't the remote file inclusion attack violate the same-origin policy?

Answer
The same origin policy is implemented in browsers, while PHP scripts are interpreted by the server. (3 points)

Problem 9. Several new DNS records are used for DNSSEC, one of them is the DS record.

a) What is in a DNS DS record?

b) When is the information used?

c) How is the information used? (3 points)

Answer

a) A hash of the DNSKEY.

b) The DS record is used when validating a DNSKEY.

c) The DS record of a DNSKEY is stored at the parent. To validate a DNSKEY, go to the parent and retrieve the corresponding DS record. Check that the hash of the DNSKEY matches the content of the DS record. To check the signature of the DS record, use the parents DNSKEY. To validate this DNSKEY, get the next parants corresponding DS record, and so on. Validation succeeds if a trusted DNSKEY is encountered along the way, typically at the top or root domain.

Problem 10. Access to directories on a web server can be defined in the *httpd.conf* file and in a *.htaccess* file.

a) Determine who has access to the directories `/www` and `/www/dir` when the following directive is given in *httpd.conf* file and *.htaccess* files located in the directories:

`httpd.conf`

```
<Directory /www>
    AllowOverride None
    Order Deny,Allow
    Deny from lth.se
</Directory>
```

`httpd.conf`

```
<Directory /www/dir>
    AllowOverride All
    Order Deny,Allow
    Deny from all
</Directory>
```

`.htaccess in /www`

```
Allow from all
```

`.htaccess in /www/dir`

```
Allow from lth.se
```

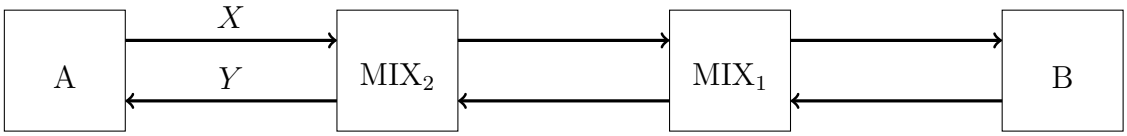
b) How is access to the two directories affected if all `Order` directives are changed to `Order Allow,Deny`?

c) Give one reason why `.htaccess` files should not be used if not absolutely needed.

Answer
a) `/www`: Everyone except computers from lth.se are allowed.
`/www/dir`: Only computers from lth.se are allowed access.
b) `/www`: Everyone is denied access.
`/www/dir`: Only computers from lth.se are allowed access.
c) They slow down access since the files are checked for each request. It can also be a

security problem since users can make changes to the server. The administrator must be careful when deciding what chages are allowed. (3 points)

Problem 11. Consider anonymous emails using two Chaum mixes, MIX_1 and MIX_2. Alice with address A sends a message M_1 to Bob with address B, including an untraceable return address. Bob replies to Alice with the message M_2.



The keys K_1 and K_2 are the public keys of MIX_1 and MIX_2 respectively. The item X prepared by Alice and sent to MIX_2 is constructed as

$$X : K_2(R_2, K_1(R_1, K_B(R_0, M_1, \underbrace{K_1(R_4, K_2(R_3, A))}_{\text{Untraceable return address}}, K_X), B))$$

a) What is K_X and what is it used for?

b) What is the purpose of R_2?

c) Determine the item Y. (5 points)

Answer
a) It is a temporary public key belonging to A, used by B to encrypt the return message M_2.
b) It adds randomness to the message encrypted by K_1 so that it is not possible to take the encrypted message part leaving MIX_2 towards MIX_1, encrypt it with K_2 and compare it with the input to MIX_2. Without R_2 it would be possible to track messages sent over MIX_2 (but not MIX_1).
c) Y is given by $R_3(R_4(K_X(R_5, M_2))), A$.

Problem 12. Digest authentication (RFC2617) calculates the digest according to

$$\text{MD5}(\ \text{MD5}(A1) : nonce : nc : cnonce : qop : \text{MD5}(A2)\ ),$$

with

$$\begin{aligned} A1 &= username : realm : password, \\ A2 &= \begin{cases} method : URI & \text{if } qop = auth, \\ method : URI : \text{MD5}(entity\text{-}body) & \text{if } qop = auth\text{-}int. \end{cases} \end{aligned}$$

a) Explain the usage and purpose of the *realm* parameter.

b) Explain the usage and purpose of the *nc* parameter?

c) Explain the usage and purpose of the *cnonce* parameter?

d) For digest authentication to work as intended, the end-user must be *aware* that the server requires digest authentication for specific pages. How can a man-in-the-middle extract username and password from an unaware user? (5 points)

Answer

a) A string explaining which password the user is expected to enter. Used as salt.

b) Nonce counter, starts at 1 and is incremented by one for every request. Prevents replay attacks.

c) A nonce that the client chooses. Prevents time-memory tradeoff attacks.

d) She can replace the digest authentication header with a basic authentication header. The user will then send username and password in cleartext (Base64-encoded).

Problem 13. Consider a Hashcash solution in which a string

$$ver : bits : date : resource : rand : counter$$

is hashed using SHA-1, where

ver is version number (currently 1),
bits indicates how costly the function is for sender,
date gives current date,
resource is recipients email address,
rand is a random number (separates different senders).

A spammer wants to send 1000 messages $M_1, \ldots, M_{1000}$ to each and every recipient on his very large mailing list, and he plans on including a Hashcash header with each mail.

a) How many calls to SHA-1 does it take to *generate* a Hashcash header with $bits = 20$? Exactly or on average?

b) How many calls to SHA-1 does it take to *verify* a Hashcash header with $bits = 20$? Exactly or on average?

c) What prevents the spammer from using the same Hashcash header when sending the message M_1 to *all* recipients on his mailing list?

d) What prevents the spammer from using the same Hashcash header when sending all messages $M_1, \ldots, M_{1000}$ to *one* specific recipient? (5 points)

Answer

a) 2^{20} times on average.

b) Exactly once.

c) Recipient's email address is included in the hashed string.

d) Nothing prevents the spammer from constructing one Hashcash header that is valid for all messages, but the mail client on the user-side typically stores previously used headers so that they cannot be used more than once.

Problem 14. Briefly explain the following terms.

a) register_globals

b) Prepared statement

c) DKIM

d) SPF

e) Perfect forward secrecy (5 points)

Answer

a) A convenience setting in PHP that automatically declares PHP variables from the corresponding GET/POST request variables.

b) A technique used to avoid SQL-injection.

c) DomainKeys Identified Mail. Associates a domain name to an email – verification that the domain has not been spoofed. Also provides integrity protection.

d) Sender Policy Framework, DNS record that states which mail servers that are allowed to send emails from a given domain.

e) The property that session keys will not be compromised if a private key is leaked at some point in the future.
