Name ..... ID ..... Section ..... Seat No .....

Sirindhorn International Institute of Technology
Thammasat University

Final Exam Answers: Semester 2, 2014

Course Title: CSS322 Security and Cryptography

Instructor: Steven Gordon

Date/Time: Friday 15 May 2015; 13:30–16:30

Instructions:

- This examination paper has 22 pages (including this page).
- Conditions of Examination: Closed book; No dictionary; Non-programmable calculator is allowed
- Students are not allowed to be out of the exam room during examination. Going to the restroom may result in score deduction.
- Turn off all communication devices (mobile phone etc.) and leave them at the front of the examination room.
- The examination paper is not allowed to be taken out of the examination room. A violation may result in score deduction.
- Write your name, student ID, section, and seat number clearly on the front page of the exam, and on any separate sheets (if they exist).

Prepared by Steven Gordon on 22 May 2015
css322y14s2e02, Steve/Courses/2014/s2/css322/assessment/final-exam.tex, r3790

Question 1 [15 marks]

Diffie-Hellman key exchange is often used in SSH:

- The server chooses public parameters prime P and base G.
- The server also selects a private S and calculates its public value $e = G^S \bmod P$.
- The server sends the three public values to the client.
- The client then chooses its private C and calculates its public f using similar equation as the server did.
- The client also calculates secret $K_C = e^C \bmod P$.
- Finally, the client sends its public value to the server, allowing the server to calculate K_S.

(a) If the server chooses $P = 29$, $G = 3$ and $S = 5$, and the client chooses $C = 6$, what is the value of K_S? Show your calculations. [3 marks]

Answer.

$$e = 3^5 \bmod 29 = 11$$

$$f = 3^6 \bmod 29 = 4$$

$$K_S = 4^5 \bmod 29 = 9$$

$$K_C = 11^6 \bmod 29 = 9$$

(b) The parameter G is normally chosen to be a primitive root of P. Explain what "primitive root of P" means. [2 marks]

Answer. *G is a primitive root of P if G to the power of any integer less than P in mod P returns a distinct value.*

(c) Prove that the server and client will always obtain the same key (even if they choose values different than those in part (a)) [2 marks].

Answer.

$$K_C = e^C \bmod P$$

$$K_C = \left(G^S \bmod P\right)^C \bmod P$$

$$K_C = \left(G^S\right)^C \bmod P$$

$$K_C = G^{SC} \bmod P$$

$$K_S = f^S \bmod P$$
$$K_S = \left(G^C \bmod P\right)^S \bmod P$$
$$K_S = \left(G^C\right)^S \bmod P$$
$$K_S = G^{CS} \bmod P$$

Therefore $K_C = K_S$.

(d) Use a diagram to illustrate a man-in-the-middle attack on Diffie-Hellman. Assume the server and client choose the same values as in part (a). Assume the attacker uses the same G and P as in part (a) and chooses a private value of 8. Show the messages sent, clearly indicating the values of parameters in those messages. Also show the value of the secrets K_C and K_S. [5 marks]

Answer.

- *Server sends* $G = 3, P = 29, e = 11$ *to client/attacker*
- *Attacker selects* $C_{att} = 8$ *and calculates* $f_{att} = 3^8 \bmod 29 = 7$
- *Attacker selects* $S_{att} = 8$ *and calculates* $e_{att} = 3^8 \bmod 29 = 7$
- *Attacker sends* $G = 3, P = 29, e_{att} = 7$ *to client*
- *Client selects* $C = 6$ *and calculates* $f = 3^6 \bmod 29 = 4$
- *Client calculates secret* $K_C = 7^6 \bmod 29 = 25$
- *Client sends* $f = 4$ *to server/attacker*
- *Attacker calculates secret* $K_{SC} = 4^8 \bmod 29 = 25$
- *Attacker calculates secret* $K_{CS} = 11^8 \bmod 29 = 16$
- *Attacker sends* $f_{att} = 7$ *to server*
- *Server calculates secret* $K_S = 7^5 \bmod 29 = 16$
- *Attacker knows secrets known by client and server*

(e) As a way to prevent the man-in-the-middle attack in SSH, the server signs the first message (containing G, P and e) using RSA. When a client first connects to the server they may see:

ssh ict.siit.tu.ac.th

```
The authenticity of host 'ict.siit.tu.ac.th' can't be established.
RSA key fingerprint is 48:9d:18:fc:9a:ef:43:8f:1d:46:96:be:53:9a:65:59.
Are you sure you want to continue connecting (yes/no)?
```

Explain what this means, including whether a man-in-the-middle attack has been prevented and why or why not. [3 marks]

Answer. *Although the server signed the message with its RSA private key, the client does not have the RSA public key of the server before the connection is established. Therefore the client cannot verify the first message; a man-in-the-middle attack may still be possible.*

Question 2 [15 marks]

Multiple choice. Circle only one answer. 1.5 marks for a correct answer. 0 marks for an incorrect answer or no answer.

(a) A cryptographic hash function, H(), that satisfies the strong collision resistance property (also called collision resistant) means it is computationally infeasible for an attacker:

i. that knows X to find Y, where H(X)=H(Y)
ii. **to find some X and Y, where H(X)=H(Y)**
iii. that knows a and X to find Y, where H(Y)=b and H(X)=a
iv. that knows a to find X, where H(X)=a
v. that knows X and Y to find a and b, where H(X)=a and H(Y)=b
vi. that knows X to find a, where H(X)=a

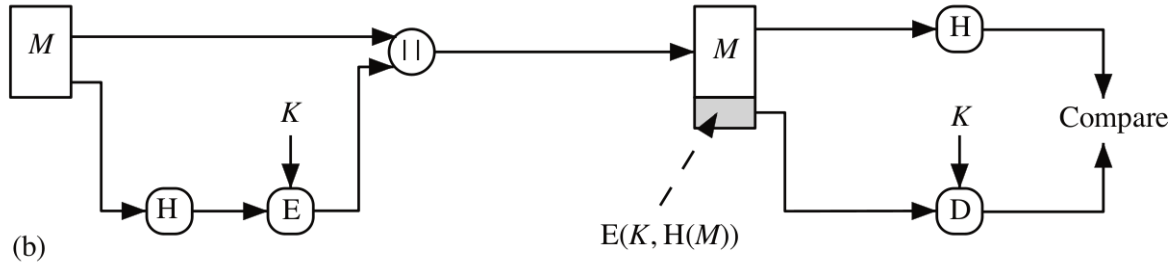
(b) A cryptographic hash function, H(), that satisfies the one way property (also called preimage resistant) means it is computationally infeasible for an attacker:

i. that knows X to find a, where H(X)=a
ii. that knows X to find Y, where H(X)=H(Y)
iii. that knows a and X to find Y, where H(Y)=b and H(X)=a
iv. to find some X and Y, where H(X)=H(Y)
v. **that knows a to find X, where H(X)=a**
vi. that knows X and Y to find a and b, where H(X)=a and H(Y)=b

(c) A cryptographically strong MAC function takes a variable sized message and a 70 bit key as input, producing a 60 bit tag as output. A brute force attack on the MAC function using a computer that can calculate MACs at a speed of 2^{40} per minute would take approximately:

i. less than 1 second
ii. more than 1 second, but less than 1 minute
iii. more than 1 minute, but less than 1 hour
iv. more than 1 hour, but less than 1 day
v. more than 1 day, but less than 6 months
vi. more than 6 months, but less than 1 year
vii. **more than 1 year, but less than 1 century**
viii. more than 1 century

(d) The scheme in Figure 4 provides:



(b)

i. confidentiality of message, integrity of message, authentication of sender

ii. authentication of sender, but no confidentiality of message and no integrity of message

iii. **authentication of sender, integrity of message, but no confidentiality of message**

iv. confidentiality of message, integrity of message, but no authentication of sender

v. integrity of message, but no confidentiality of message and no authentication of sender

vi. no confidentiality of message, no integrity of message, no authentication of sender

(e) User A sends a message, M, and corresponding tag, T, to user B. The tag is calculated as: $T = \text{MAC}(K, M)$. A malicious user, C, intercepts the message before it reaches B. User C modifies the message and forwards the modified message with the unchanged tag to B. User B verifies the received message. Assuming a cryptographically strong MAC function is used, what is the reason that the attack by C will be unsuccessful?

i. MAC of the same message using the same key will always produce identical tags

ii. MAC of the same message but using two different keys should produce different tags

iii. MAC of two different messages using the same key will always produce identical tags

iv. MAC of the same message but using two different keys will always produce identical tags

v. MAC of the same message using the same key should produce different tags

vi. **MAC of two different messages using the same key should produce different tags**

(f) A malicious user C, masquerading as user A, sends a message, M, and corresponding tag, T, to user B. The tag is calculated as: $T = \text{MAC}(K, M)$. User B verifies the received message. Assuming a cryptographically strong MAC function is used, what is the reason that the attack by C will be unsuccessful?

i. MAC of the same message but using two different keys will always produce identical tags

ii. **MAC of the same message but using two different keys should produce different tags**

iii. MAC of the same message using the same key should produce different tags

iv. MAC of the same message using the same key will always produce identical tags

v. MAC of two different messages using the same key should produce different tags

vi. MAC of two different messages using the same key will always produce identical tags

(g) HMAC is best described as an algorithm that:

i. **converts a hash function into a MAC function**

ii. converts a hash function into a public key crypto function

iii. converts a public key crypto function into a MAC function

iv. converts a public key crypto function into a hash function

v. converts a MAC function into a hash function

vi. converts a MAC function into a public key crypto function

(h) Consider a hash function H() that works only on a fixed length input, i.e. all messages are 200 bits. The function produces random hash values as output, each value is 128 bits. In theory, the average number of messages that collide on the same output is:

i. 2^{328}

ii. 2^{200}

iii. 2^{128}

iv. 2^{72} *correct*

v. 328

vi. 200

vii. 128

viii. 72

ix. 1

x. 0

(i) Which equation best represents how user A signs message M to be sent to user B?

i. $M||E(PU_A, H(M))$
ii. $M||E(PR_A, H(M))$ *correct*
iii. $M||E(PU_B, H(M))$
iv. $M||E(PR_B, H(M))$
v. $M||E(PU_A, H(M||S))$
vi. $M||E(PR_A, H(M||S))$
vii. $M||E(PU_B, H(M||S))$
viii. $M||E(PR_B, H(M||S))$
ix. $M||E(PU_A, H(M||PU_A))$
x. $M||E(PR_A, H(M||PR_A))$
xi. $M||E(PU_B, H(M||PU_B))$
xii. $M||E(PR_B, H(M||PR_B))$

(j) The security of Diffie-Hellman key exchange depends on which problem being difficult to solve:

i. prime factorization
ii. calculating Euler's totient
iii. generating true random numbers
iv. **discrete logarithm**
v. modular arithmetic
vi. collision resistance

Question 3 [7 marks]

Consider the two mechanisms illustrated in Figures 1 and 2.

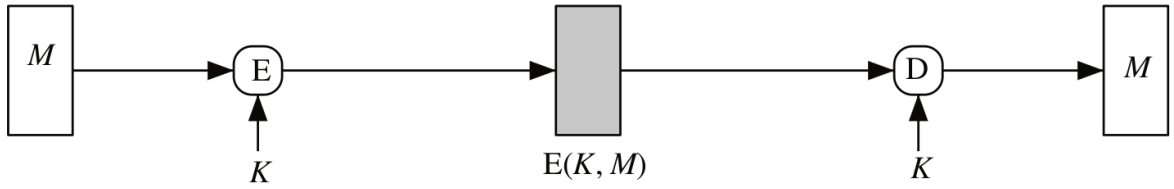


Figure 1: Security mechanism 1

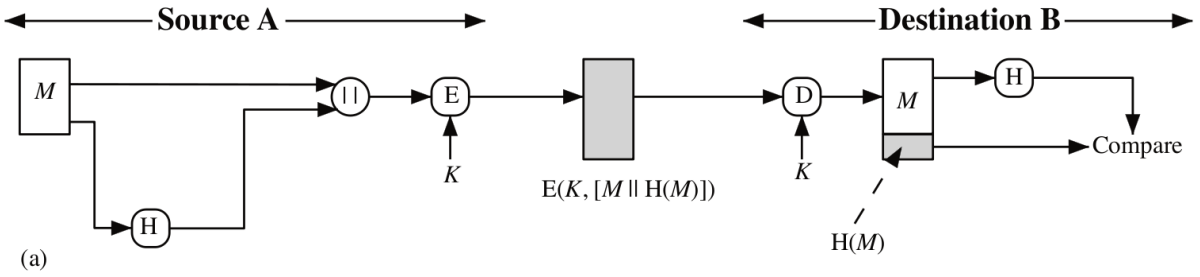


Figure 2: Security mechanism 2

(a) Assuming the message M being sent is a randomly generated cryptographic key, explain why mechanism 2 (Figure 2) can be used for authentication, but mechanism 1 (Figure 1) cannot. [3 marks]

Answer. *If the message is random then when the receiver decrypts it has know way to determine if the plaintext obtained is correct or not. Therefore the receiver cannot be sure that the correct key has been used or if the message has been modified. By using a hash, even the message is random, the hash of the received message should match the hash received.*

(b) Considering only mechanism 2 (Figure 2) for any type of message, if the hash function is not weak collision (second pre-image) resistant, then is it possible for an attacker to defeat the authentication mechanism? If yes, explain how. If no, explain why not. [2 marks]

Answer. *No. Since the hash value is encrypted, there is no way for the attacker to know which hash value to search for.*

(c) If in mechanism 2 the algorithms are AES128 and SHA512, and the message M is 6400 bits long, assuming no padding or packet headers are needed/used, how many bits are sent between A and B? [2 marks]

Answer. *The 6400 bit message is concatenated with the 512 bit hash value to give 6,912 bits. This is encrypted using a 128 bit key, producing 6,912 bits of ciphertext to be sent from A to B.*

Question 4 [13 marks]

A (digital) certificate contains a users public key (PU), identity (ID), a timestamp (T), as well as those three values signed by a trusted entity. Consider a system where the trust relationships between users are shown in a hierarchy as in Figure 3 where a user trusts the user on the next level up in the hierarchy. For example, users A and B trust user X; user X trusts user Z.

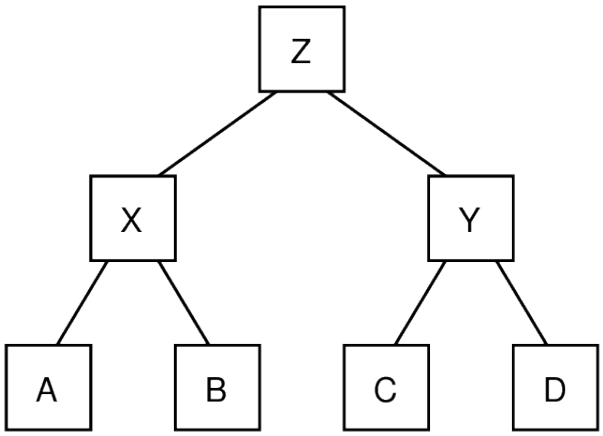


Figure 3: Trust relationship

In the following questions you may use the operators: $||$ for concatenation, $E(k, p)$ and $D(k, c)$ for encryption and decryption, and $H()$ for a hash function.

(a) Write an equation that shows the certificate of user A, i.e. $C_A = \ldots$ [3 marks]

Answer.

$$C_A = ID_A||PU_A||T||E(PR_X, H(ID_A||PU_A||T))$$

(b) Who signs C_Z? [1 mark]

Answer. *C_Z is a self-signed certificate; Z signs its own certificate*

Assume all users except user B already have their own certificate and the certificate of the other user they trust. For example, A has C_A and C_X. Consider user B.

(c) What algorithm can B use to generate its own key pair? [1 mark]

Answer. *RSA*

(d) Which user does B send a Certificate Signing Request to? [1 mark]

Answer. X

All users, including B, now have their own certificate and the certificate of the other user they trust. Consider A wanting to communicate confidentially with B.

(e) Draw a diagram that shows the steps that A and B (and others if necessary) take to exchange public keys. The diagram should be similar to those seen in lecture (e.g. like Figure 5 or 4), labelling the messages with numbers to indicate the order and showing the contents of messages. [2 marks]

Answer. *A sends C_A to B and B sends C_B to A*

(f) Explain how A verifies the information it receives from B, including what information A must know to perform the verification. [2 marks]

Answer. *A needs C_X. A uses PU_X from C_X to decrypt the signature of C_B. If the decrypted value matches the hash of the received public key then it is successfully verified.*

Now consider B wanting to communicate confidentially with C. They exchange their public keys.

(g) Explain how C verifies the information it receives from B, including what information C must know to perform the verification and what other communications may need to take place to complete the verification. [3 marks]

Answer. *When C receives the certificate of B (which is signed by X), C needs the certificate of X (i.e. the public key) to verify. And C must trust that it is inded the certificate of X. Initially C trusts Y, meaning C has Y's certificate. Y obtains the certificate of X, and since the certificate of X is signed by Z, and Y trusts Z (i.e. has the certificate of Z), Y verifies the certificate of X. Now Y can sign the public key of X and send to C. Since C trusts Y, C now has the certificate of X and can verify the certificate of B.*

Question 5 [6 marks]

The following are a selection of Linux commands used for cryptographic operations. Some commands have selected parts hidden with XXX.

(a) `openssl genpkey -algorithm XXX -out privkey.pem`

(b) `openssl pkey -in privkey.pem -out XXX -pubout`

(c) `openssl genpkey -genparam -algorithm DH -out dhp.pem`

(d) `openssl pkeyutl -derive -inkey key1.pem -peerkey pub2.pem -out secret.bin`

(e) `openssl req -new -key privkey.pem -out XXX`

(f) `xxd -l 32 -g 32 -c 32 secret.bin`

(g) `openssl rand 16 -hex`

(h) `openssl s_client -connect XXX:443`

Select the most appropriate command from above that is used to perform each of the following operations. To answer, in the space for each operation, give the letter, from between *a* and *g*, of the command. [1.5 marks each]

(a) Create a request message to be sent to a CA so that CA can generate a X.509 certificate *e*

(b) Generate global public Diffie-Hellman parameters *c*

(c) Test a secure web connection *h*

(d) Create a secret using information from another entity in a key exchange *d*

Question 6 [7 marks]

(a) Explain how key-based (also called password-less) authentication works in SSH. In your explanation, include the information that must be created/exchanged to configure key-based authentication, and how the authentication is perform when you "SSH into a server". (You don't need to list specific files/directories, just the concepts). [3 marks]

Answer. *The client generates a public/private key pair. The public key must be transferred to the server in advance. When the client SSH's into the server the client signs a message with its private key. The server verifies it is the client using the clients public key.*

You setup your web browser on your computer to use a SOCKS Proxy on your localhost. This sends everything that your browser normally would send across your Internet connection instead to a SOCKS proxy server on your computer. You then configure a SSH client on your computer to act as a SOCKS proxy server and to connect to the SSH server ict.siit.tu.ac.th, using the command:

```
ssh -N -D 12345 ict.siit.tu.ac.th
```

(b) Assuming you are using your computer (laptop) in the above configuration while on holiday in Australia. You visit the website http://sandilands.info/ (hosted in Japan). Draw a diagram that shows the flow from browser to web server. On your diagram you must clearly show what parts of the data flow are encrypted. [2.5 marks]

Answer. *Browser (AU) --encrypted--> ICT (TH) ------> Server (JP)*

(c) In the configuration in part (b), what are the port numbers used on the servers at the following locations: [1.5 marks]

i. Australia:

ii. Japan:

iii. Thailand:

Answer. *12345 (SOCKS) in Australia, 22 (SSH) in Thailand, 80 (HTTP) in Japan*

Question 7 [12 marks]

Consider a system with 100 users (e.g. user A, user B, ... user Z, user AA, user AB, ...). Confidentiality of communications between users must be provided using symmetric key cryptography. Figures 4 and 5 show two alternative protocols for key distribution in the system for an example when user A wants to communicate with user B. First consider the protocol in Figure 4.

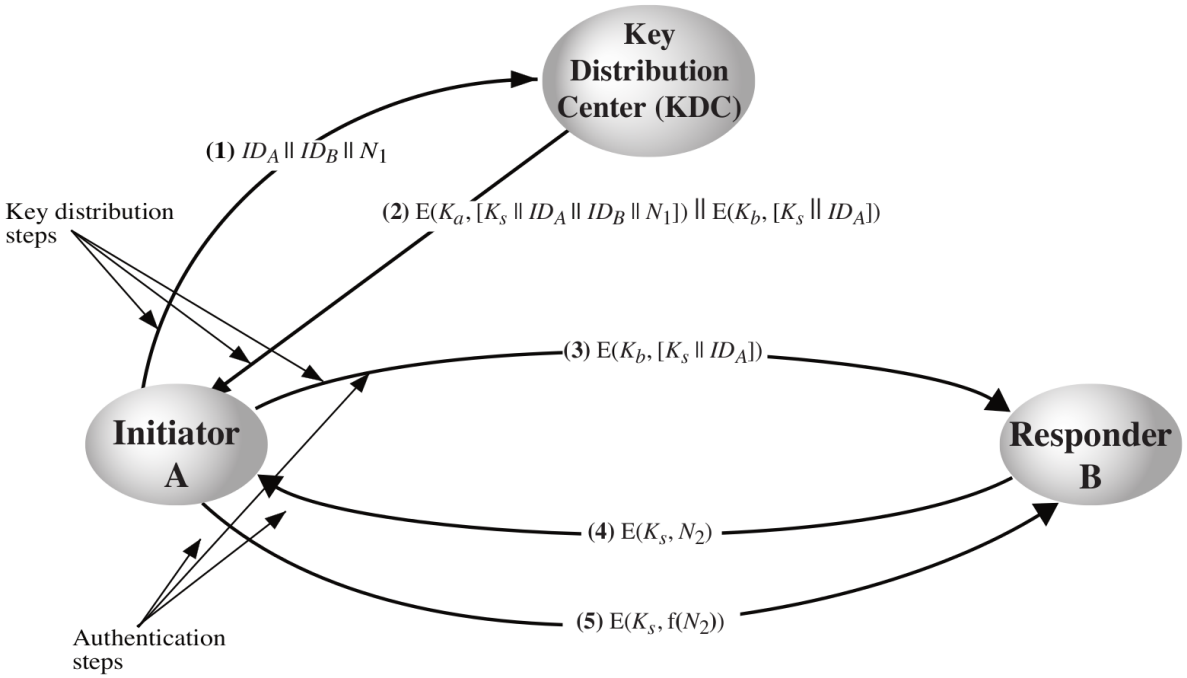


Figure 4: Key distribution protocol 1

(a) What is the set of keys that is assumed to be known by each entity *before* the protocol is applied? [2 marks]

Answer. *User A knows K_a; user B knows K_b; ...; user Z knows K_z; and KDC knows K_a, K_b, ..., K_z*

(b) What is the set of additional keys that are known by each entity *after* the protocol is applied? (that is, in addition to the keys known in part (a)) [2 marks]

Answer. *User A also K_s; user B also knows K_s; and KDC also knows K_s*

(c) If an attacker intercepts all five messages during the protocol operation, list all the items that the attacker will know. [1 mark]

Answer. ID_A, ID_B, N_1

(d) If after the protocol operation (i.e. all five messages are sent) an attacker later replays message (3), explain how the replay attack would be detected. [2 marks]

Answer. *User B responds with message (4), containing a random nonce encrypted with K_s. B is then expecting message (5) in return (if it does not receive it or receives it with the wrong nonce, then the attack is detected). If the malicious user intercepts message (4) it cannot determine N_2 because it doesn't know K_s, therefore B will not receive the expected response (attack detected). If user A receives message (4) then the attack is detected because A wasn't expecting this message since A did not send message (3).*

Now compare the protocol in Figure 4 with the protocol in Figure 5. Both are used to automatically share session keys between the 100 users.

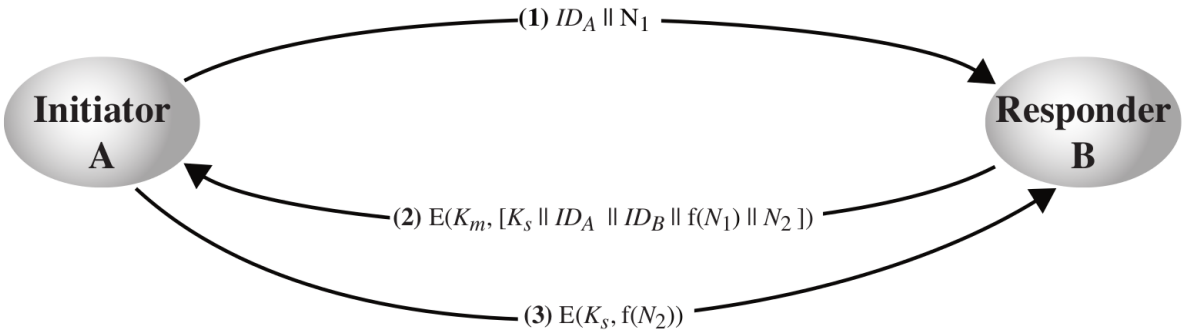


Figure 5: Key distribution protocol 2

(e) How many master keys needed in the system using the scheme from Figure 4? [1.5 marks]

Answer. *100*

(f) How many master keys needed in the system using the scheme from Figure 5? [1.5 marks]

Answer. *4950*

(g) Explain an advantage of the protocol in Figure 4 compared to that in Figure 5? [1 mark]

Answer. *Fewer keys to be manually distributed before the protocol operation.*

(h) One advantage of using the protocol in Figure 5 (compared to that in Figure 4) is that it avoids performance bottlenecks at KDC. Explain another advantage of Figure 5. [1 mark]

Answer. *No need to trust KDC*

Question 8 [8 marks]

A company has developed a new protocol, called *BAHTP*, that is used by a client application on computers in shops around Bangkok to send sales information to a central server in the company main office in Rangsit. The protocol uses TCP/IP. Based on your expert knowledge of OpenSSL libraries, you have been hired by the company to modify the client/server applications so that all communications between them are secure.

(a) Draw a protocol stack of a computer using Ethernet physical and data link layers, that illustrates the protocols in use by the secure client application. [2 marks]

Answer.

BAHTP

SSL/TLS

TCP

IP

Ethernet DLL

Ethernet PHY

When using the secure application, a secure session and connection has been established. The following information is stored by the client computer for this session/connection.

- Session ID: id
- Compression method: null
- CipherSuite: TLS_DHE_RSA_WITH_AES_128_CBC_SHA
- Master secret: s
- Server random: r_s
- Client random: r_c
- Server MAC secret: m_s
- Client MAC secret: m_c
- Server encrypt key: e_s
- Client encrypt key: e_c

Figure 6 shows the general operation of SSL record protocol.

(b) To generate the master secret s, a premaster secret is used. What algorithm is used for the client and server to share a premaster secret. [1 mark]

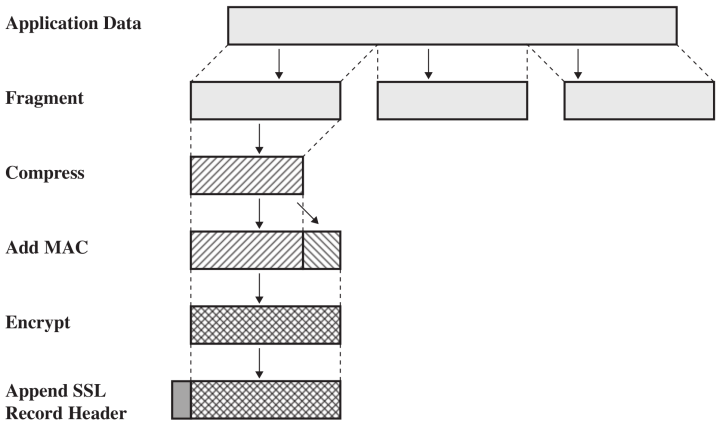


Figure 6: SSL Record Protocol Operation

Answer. *Diffie-Hellman (Ephemeral)*

(c) What algorithm/cipher is used for the client the authenticate the server? [1 mark]

Answer. *RSA*

(d) The master secret s is used, in combination with the random values, identities of the client/server and hash functions, to create multiple keys. Explain a security advantage of having multiple secrets/keys. [1 mark]

Answer. *In the above there is a master secret, as well as separate encrypt and MAC keys. An advantage is that the master secret is used very few times to encrypt data sent; instead the encrypt keys are, which may be changed regularly. Therefore an attacker has limited time to try to discover a key. Also, if one encrypt key is compromised, then other data may still be protected.*

(e) Write an equation that expresses the SSL record operation on a single fragment, F from the client application that produces the packet to be sent P. Use the variables above and || for the concatenate/append operator. For function names you *must* use the algorithm names (i.e. you cannot use E() for encrypt, H() for hash; refer to specific algorithms). Denote the SSL header as SSL. [3 marks]

Answer.

$$P = SSL||\text{AES\_128\_CBC}(e_c, F||\text{HMACSHA}(m_c, F))$$

Question 9 [7 marks]

Consider the key distribution scheme in Figure 7.

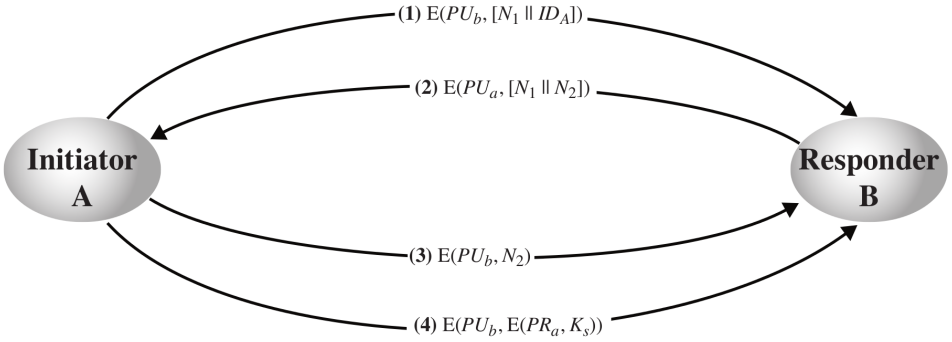


Figure 7: Key distribution scheme

(a) What keys are assumed to be known by B before the scheme starts? [1 mark]

Answer. PU_b, PR_b, PU_a

(b) What is the purpose of this scheme? What are they distributing? [1 mark]

Answer. *Distribute* K_s

(c) What is N_1? Describe it or given an example of how it is selected. [1 mark]

Answer. *Nonce. A number used only once, e.g. random number*

(d) Consider an malicious user M that can intercept/modify any messages between A and B. Draw a diagram to illustrate M performing a man-in-the-middle attack, where message (1) is intercepted, and a new message sent to B containing ID_M. Explain why/when the attack will be detected. [4 marks]

Answer.

- *Message (1) is received by the malicious user, but it cannot be decrypted (since malicious user doesn't have* PR_b
- *Malicious user creates a new message:* $E(PU_b, [N_3||ID_M])$. *Note that the malicious user includes their ID so that later B will use* PU_M *to encrypt. Also the nonce included,* N_3*, is different from* N_1 *(since* N_1 *was encrypted).*
- *B responds with:* $E(PU_M, [N_3||N_2])$
- *Malicious user intercepts and decrypts to learn* N_2
- *Malicious user sends to A:* $E(PU_a, [N_3||N_2])$
- *When A decrypts the message it realizes* N_3 *is different than* N_1*, therefore detecting the attack*