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

Sirindhorn International Institute of Technology
Thammasat University

Final Exam Answers: Semester 2, 2012

Course Title: CSS322 Security and Cryptography

Instructor: Steven Gordon

Date/Time: Friday 1 March 2013; 13:30–16:30

Instructions:

- This examination paper has 18 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.
- 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).
- The examination paper is not allowed to be taken out of the examination room. A violation may result in score deduction.

Security and Cryptography, Semester 2, 2012

Prepared by Steven Gordon on 18 February 2013
CSS322Y12S2E02, Steve/Courses/2012/s2/css322/assessment/final-exam.tex, r2693

Question 1 [12 marks]

The encryption algorithm of RSA is defined as:

$$C = M^e \bmod n$$

(a) What is the decryption algorithm of RSA? [1 mark]

Answer.

$$M = C^d \bmod n$$

(b) What is the public key in RSA? [1 mark]

Answer. *e, n*

(c) What is the private key in RSA? [1 mark]

Answer. *d, n*

(d) Describe the steps for generating the public/private key pair. You must state the conditions/properties of any values to be selected or calculated. (You do not need to explain why those conditions are necessary) [3 marks]

Answer.

- *Select two large prime integers, p and q.*
- *Calculate $n = p \times q$ and $\phi(n) = (p-1) \times (q-1)$.*
- *Select e such that it is relatively prime with $\phi(n)$.*
- *Calculate d, the multiplicative inverse of e in mod $\phi(n)$.*

Based on the definition of RSA, there are three theoretical approaches for an attacker, knowing only public information, to discover the private information and/or a plaintext message.

(e) What public information is it assumed that an attacker knows in RSA? (Refer to the variables defined in parts (a) to (d)). [1 mark]

Answer. *Attacker knows: e, n, C*

(f) Describe one of the three theoretical approaches that an attacker can use. [3 marks]

Answer.

- *Approach 1. Determine p and q by factoring n into its prime factors, so thatcan be easily calculated, and subsequently d.*
- *Approach 2. Given C, e and n, calculate the inverse of $C = M^e \bmod n$. That is, find an M such that: $e = dlog_{M,n}(C)$.*
- *Approach 3. From n, calculate without knowing p and q.*

(g) What makes the above approach practically impossible for an attacker to use? [2 marks]

Answer.

- *Approach 1. Determining the prime factors of a large number is computationally hard.*
- *Approach 2. Calculating the discrete logarithm (inverse exponential) for large numbers is computationally hard.*
- *Approach 3. Calculating for large n is computationally hard.*

Question 2 [6 marks]

(a) User A wants to digitally sign a document M and send it to B. Give a function that describes how the signing is performed (you must also describe all variables used) and explain what is sent from A to B. [2 marks]

Answer. *Signature = E(PRA,M) where PRA is the private key of A The signature and M are sent to B (e.g. concatenated together).*

(b) User A wants to send a MAC authenticated message M to B. Give a function that describes how the authentication data is generated (you must also describe all variables used) and explain what is sent from A to B. [2 marks]

Answer. *MACdata = MAC(S,M) where S is a shared secret key with B. The MAC data and M are sent to B (e.g. concatenated together).*

(c) Explain why MAC-based authentication cannot be used as a digital signature. [2 marks]

Answer. *A MAC function uses a shared secret key. A message authenticated with a MAC function confirms that the message was generated by either the of the parties that has the secret. It does not confirm which of the two parties generated the message (which is the purpose of a digital signature).*

Question 3 [5 marks]

Consider the protocol in Figure 1.

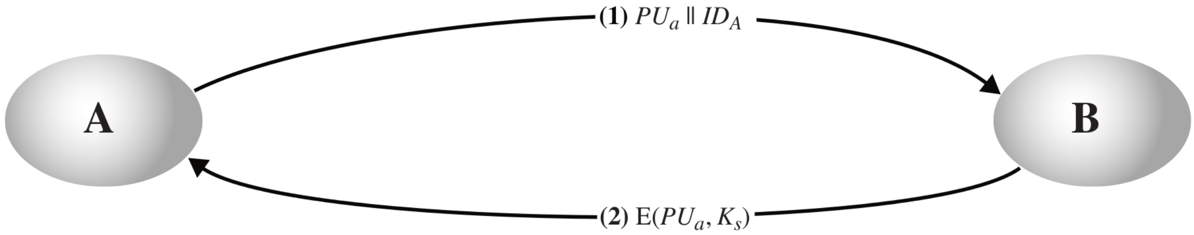

Figure 1: Protocol

(a) Explain the purpose of this protocol. That is, what is the objective of performing this steps? [2 marks]

Answer. *The purpose is to exchange a secret key, i.e. for both A and B to know the secret K_s.*

(b) Draw a diagram that shows how user C can perform a man-in-the-middle attack when this protocol is used? [3 marks]

Answer. *The diagram should show the following messages:*

$$A \rightarrow C : PU_a || ID_A$$

$$C \rightarrow B : PU_c || ID_A$$

$$B \rightarrow C : E(PU_c, K_s)$$

$$C \rightarrow A : E(PU_a, K_s)$$

Question 4 [7 marks]

Consider the X.509 certificate in Listing 1.

Listing 1: X.509 Certificate

```
Certificate:
    Data:
        Version: 3 (0x2)
        Serial Number: 3 (0x3)
        Signature Algorithm: sha1WithRSAEncryption
        Issuer: C=TH, ST=Pathumthani, O=TrustUs, OU=Crypto,
                CN=TrustUsCrypto/emailAddress=security@trustus.co.th
        Validity
            Not Before: Jan 25 02:25:10 2011 GMT
            Not After : Jan 25 02:25:10 2012 GMT
        Subject: C=TH, ST=Pathumthani, O=TheAuthorityCompany,
                 CN=TheAuthorityCompany/emailAddress=crypto@auth.co.th
        Subject Public Key Info:
            Public Key Algorithm: rsaEncryption
            RSA Public Key: (1024 bit)
                Modulus (1024 bit):
                    00:aa:1f:cf:01:2f:d3:2e:80:63:98:1b:0f:16:5d:
                    dd:af:e2:38:de:78:88:56:b6:14:2b:61:79:92:0b:
                    f3:7f:b6:89:7b:d0:fc:59:5a:1a:be:24:61:39:d5:
                    4d:80:3a:40:2b:7c:89:ef:5e:50:a5:3b:44:68:a9:
                    7f:97:d9:c4:9a:bf:b6:97:eb:4c:87:0d:00:96:b4:
                    f9:ea:8c:6a:cb:e0:bd:f8:a8:1f:82:d3:2b:23:3c:
                    b6:54:85:37:5b:13:1a:2e:be:0d:20:52:c5:98:b6:
                    4c:97:67:6e:b2:43:04:3f:01:41:8e:e0:2f:38:1f:
                    e1:cc:cf:0d:c2:5f:0a:d3:e1
                Exponent: 65537 (0x10001)
        X509v3 extensions:
            X509v3 Basic Constraints:
                CA:FALSE
            Netscape Comment:
                OpenSSL Generated Certificate
            X509v3 Subject Key Identifier:
                EA:1C:DC:C5:16:F2:9D:BC:61:5E:A8:D2:67:2A:06:13:C5:64:8A:AE
            X509v3 Authority Key Identifier:
                keyid:61:52:40:EA:7F:E0:EC:77:41:F6:4F:6F:7C:49:EB:05:C1:56:6D:49

    Signature Algorithm: sha1WithRSAEncryption
        a5:7a:36:91:ef:11:46:58:74:37:87:81:7a:99:ff:b6:40:4a:
        80:6a:07:69:e3:3c:33:9a:fd:31:50:e9:9f:bf:ff:36:a4:34:
        21:50:49:70:e0:88:b3:01:c9:51:26:8b:1e:8b:34:ca:4c:3c:
        a2:ab:0a:a3:b3:39:c0:fb:88:72:98:69:c9:af:42:b2:48:1b:
        4e:4a:76:e8:b4:c7:d4:f8:15:d2:5e:f8:69:fc:53:0c:ca:85:
        84:ea:e5:36:17:20:65:fc:d0:3e:d1:33:17:f7:d1:40:f8:3d:
        2a:87:f8:3c:66:8e:43:62:ea:02:ef:7a:d4:a7:55:e9:d9:2d:
        38:1a
```

```
-----BEGIN CERTIFICATE-----
MIIC5zCCAlCgAwIBAgIBAzANBgkqhkiG9w0BAQUFADCBnzELMAkGA1UEBhMCVEgx
GDASBgNVBAgTC1BhdGh1bXRoYW5pMREwDwYDVQQHEwhCYW5na2FkaTENMAsGA1UE
ChMEU0lJVDEMMAoGA1UECxMDSUNUMR4wHAYDVQQDExVDZXJ0aWZpY2F0ZSBBdXRo
b3JpdHkxKjAoBgkqhkiG9w0BCQEWG2NzczMyMi1jYUBpY3Quc2lpdC50dS5hYy50
aDAeFw0xMTAxMjUwMjI1MTBaFw0xMjAxMjUwMjI1MTBaMFYxCzAJBgNVBAYTAlRI
MRQwEgYDVQQIEwtQYXRodW10aGFuaTENMAsGA1UEChMEU0lJVDEMMAoGA1UECxMD
SUNUMRQwEgYDVQQDEwtEZW1vIFVzZXIgMjCBnzANBgkqhkiG9w0BAQEFAAOBjQAw
gYkCgYEAqh/PAS/TLoBjmBsPFl3dr+I43niIVrYUK2F5kgvzf7aJe9D8WVoaviRh
OdVNgDpAK3yJ715QpTtEaKl/l9nEmr+2l+tMhw0AlrT56oxqy+C9+KgfgtMrIzy2
VIU3WxMaLr4NIFLFmLZMl2duskMEPwFBjuAvOB/hzM8Nwl8KBKMCAwEAAaN7MHkw
CQYDVR0TBAIwADAsBglghkgBhvhCAQ0EHxYdT3BlblNTTCBHZW5lcmF0ZWQgQ2Vy
dGlmaWNhdGUwHQYDVR0OBBYEFOoc3MUW8p28YV6o0mcqBhPFZIquMB8GA1UdIwQY
MBaAFGFSQOp/4Ox3QfZPb3xJ6wXBVm1JMA0GCSqGSIb3DQEBBQUAA4GBAKV6NpHv
EUZYdDeHgXqZ/7ZASoBqB2njPDOa/TFQ6Z+//zakNCFQSXDgiLMByVEmix6LNMpM
PKKrCqOzOcD7iHKYacmvQrJIG05Kdui0x9T4FdJe+Gn8UwzKhYTq5TYXIGX80D7R
Mxf30UD4PSqH+DxmjkNi6gLvetSnVenZLTga
-----END CERTIFICATE-----
```

(a) Whose certificate is this? [1 mark]

Answer. *The user* The Authority Company

(b) Whose RSA key is included in the certificate? [1 mark]

Answer. *The user* The Authority Company

(c) What are the last two hexadecimal digits of e in the users RSA key? [1 mark]

Answer. *The exponent, e, is 65537 in decimal, or 10001 in hex. The answer is 01.*

(d) What are the last two hexadecimal digits of n in the users RSA key? [1 mark]

Answer. *The modulus, n, is given in hex and ends with e1.*

In general, an X.509 certificate for user A can be expressed as:

$$C_A = Data||S$$

where $Data$ is the concatenation of the fields: Version, SerialNumber, SignatureAlgorithm, Issuer, Validity, Subject, SubjectPublicKeyInfo and X509v3extensions.

(e) Write an equation for how S is calculated in the certificate in Listing 1? You must use the names of algorithms used in the above certificate (i.e. you cannot use E()), as well as clearly identify which user each key belongs to. You may use the variable $Data$ in your equation to represent the concatenation of various fields. [3 marks]

Answer.

$$S = RSA(PR_{TrustUsCrypto}, SHA1(Data))$$

Question 5 [8 marks]

The original standard for encryption in a wireless LAN (WiFi) is called Wired Equivalent Privacy (WEP). Early devices that used WEP allowed the user to select a 10 hexadecimal digit value, which was combined with a 24-bit initialisation vector to produce the encryption key. The IV was sent as plaintext and changed for every packet sent.

(a) What is the entropy of the user selected value? [2 marks]

Answer. *10 hexadecimal digits is 40 bits. Therefore the entropy is 40*

(b) An alternative to entering a hexadecimal value would be to allow the user to enter the key using the set of lowercase English letters as well as numbers. How many characters are needed for the key entered using letters and numbers? [2 marks]

Answer. *With 36 possible characters, there are 5.17 bits per character. So would need at least 8 characters to create 40 bits.*

When a user can select a string from letters and numbers they normally do not chose it randomly. One study has calculated the approximate entropy of such strings if the user can choose: any value; any value, except for those in a dictionary. The entropy values for different length strings is shown in Table 1.

Table 1: Entropy of user chosen ASCII strings

Length	*Any Value*	*Any Value, except Dict.*
6	12	20
10	19	29
14	25	34
18	31	37
20	34	39
22	36	41
24	38	43
30	44	49
40	54	59

An improved security protocol for wireless LAN is called WiFi Protected Access (WPA). It allows a 256-bit key, generated from a password chosen by the user of between 8 to 63 characters (from the set of lowercase letters and numbers). Assume a malicious user can attempt to guess the password at a rate of 1,000 guesses per second.

(c) If the user chose a 14 character password and was allowed any value, on average approximately how long would it take the malicious user to guess the password? [2 marks]

Answer. *From the table, a 14 character password allowing any value gives a password with an entropy of 25. That gives* 2^{25} *possible passwords. On average half of the passwords need to be guessed giving* $2^{24} = 16777216$ *guesses. At 1000 per second it takes the attacker 16777.216 seconds.*

(d) If the user is allowed to choose a password with any value, except that from a dictionary, then what is the minimum password length that offers the same strength as the 10 hexadecimal digit value in part (a)? [2 marks]

Answer. *From part (a) we need a password with entropy of 40. From the table the password must be 22 characters to give an entropy of 40.*

Question 6 [11 marks]

In Diffie-Hellman key exchange, user Supree can calculate his public value S as:

$$S = a^{X_S} \bmod n$$

where $X_S < n$, n is a prime number, a is a primitive root of n and $a < n$. Assume Supree wants to exchange a secret, K, with user Usa.

(a) What is the equation for Usa to calculate her public value, U? [1 mark]

Answer. $U = a^{X_F} \bmod n$

(b) What value does Usa send to Supree in the Diffie-Hellman exchange? [1 mark]

Answer. U

(c) What is the equation for Supree to calculate the secret, K_s? [2 marks]

Answer. $K_s = U^{X_s} \bmod n$

(d) What value(s) are public in this Diffie-Hellman exchange (that is, assumed that a malicious user knows them)? [2 marks]

Answer. *a, n, U, S*

(e) What value(s) should only be known by Usa (that is, no other users should know them)? [2 marks]

Answer. X_U

(f) Prove that the secret calculated by Usa, K_U, is the same as the secret calculated by Supree, K_S. Show the detailed steps of your proof. [3 marks]

Answer.

$$\begin{aligned} K_U &= S^{X_U} \bmod n \\ &= (a^{X_S} \bmod n)^{X_U} \bmod n \\ &= (a^{X_S})^{X_U} \bmod n \\ &= a^{X_S X_U} \bmod n \end{aligned}$$

$$\begin{aligned} K_S &= U^{X_S} \bmod n \\ &= (a^{X_U} \bmod n)^{X_S} \bmod n \\ &= (a^{X_U})^{X_S} \bmod n \\ &= a^{X_U X_S} \bmod n \end{aligned}$$

Therefore $K_U = K_S$.

Question 7 [11 marks]

Consider a system with 26 users (e.g. user A, user B, ... user Z). Confidentiality of communications between users must be provided using symmetric key cryptography. Figures 2 and 3 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 2.

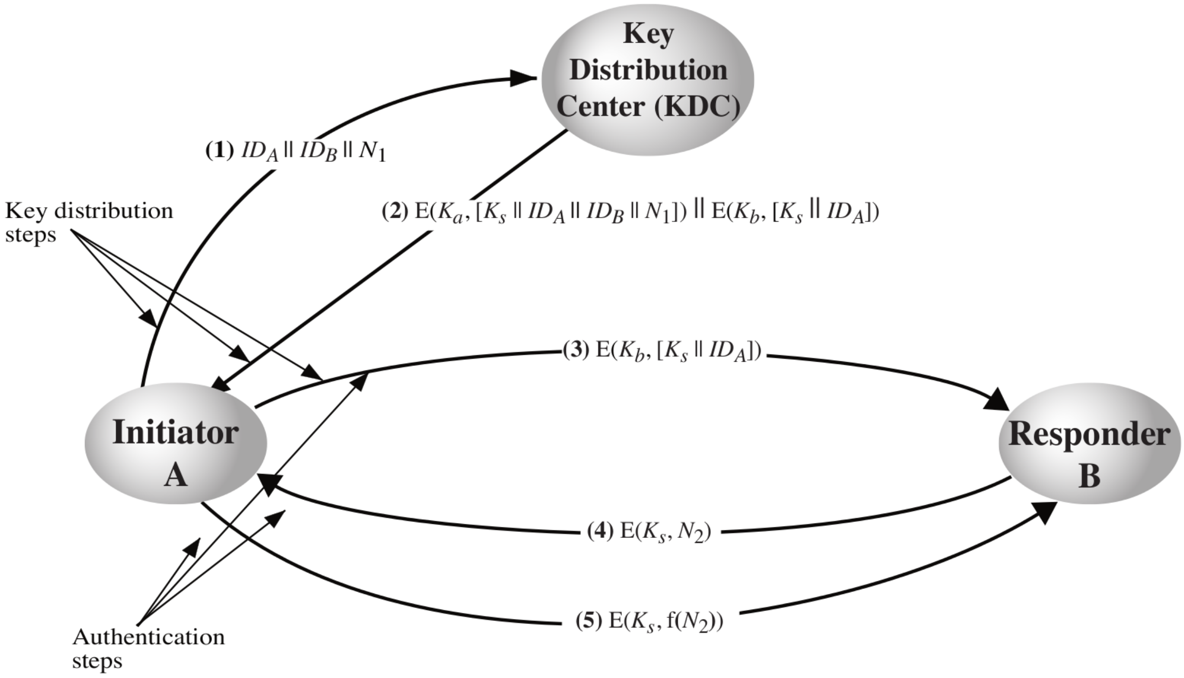
Key Distribution Center (KDC)
(1) $ID_A \| ID_B \| N_1$
Key distribution steps
(2) $E(K_a, [K_s \| ID_A \| ID_B \| N_1]) \| E(K_b, [K_s \| ID_A])$
(3) $E(K_b, [K_s \| ID_A])$
Initiator A
Responder B
(4) $E(K_s, N_2)$
(5) $E(K_s, f(N_2))$
Authentication steps

Figure 2: 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 2 with the protocol in Figure 3.

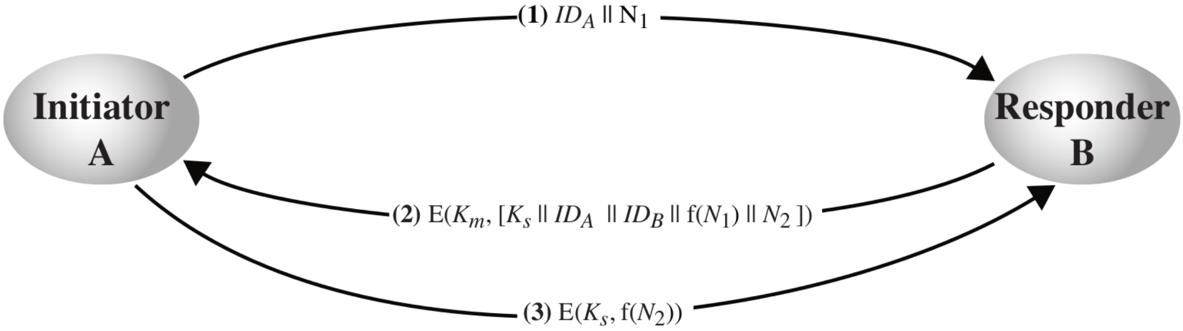

Figure 3: Key distribution protocol 2

(e) What is the total number of keys that user A is assumed to know *before* the protocol is applied in Figure 3? [2 marks]

Answer. *User A must share master keys with all other users, i.e. 25*

(f) Explain an advantage of the protocol in Figure 2 compared to that in Figure 3? [1 mark]

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

(g) One advantage of using the protocol in Figure 3 (compared to that in Figure 2) is that it avoids performance bottlenecks at KDC. Explain another advantage of Figure 3. [1 mark]

Answer. *No need to trust KDC*

Question 8 [12 marks]

Consider the mechanism illustrated in Figure 4.

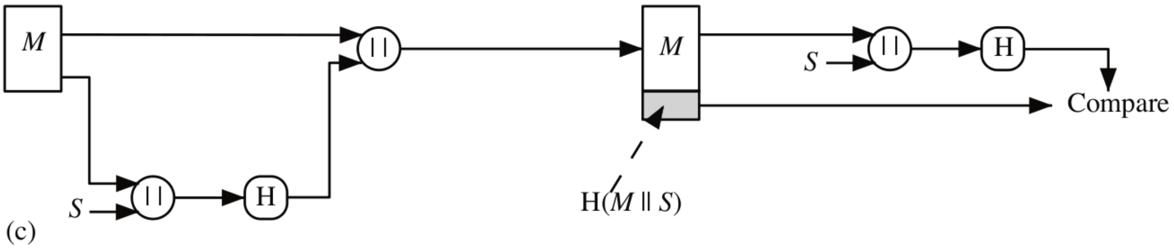

Figure 4: Security mechanism 1

(a) What is a security service that this mechanism provides? [1 mark]

Answer. *Authentication, data integrity*

(b) Explain (or define) the *one-way property* (also called *pre-image resistant property*) of a hash function. [2 marks]

Answer. *Computationally hard to determine the input of a hash function, given only the hash function and the output hash value*

(c) Explain how an attacker can defeat the above security service if the function H() did not have the one-way property. [2 marks]

Answer. *If the one-way property does not hold, then from the hash value,* $\mathrm{H}(M||S)$ *the attacker can find* $M||S$*. Since the attacker also knows* M *they can find* S*, the shared secret. Once they know the secret they could send a message to B, pretending to be A.*

Consider the mechanism illustrated in Figure 5

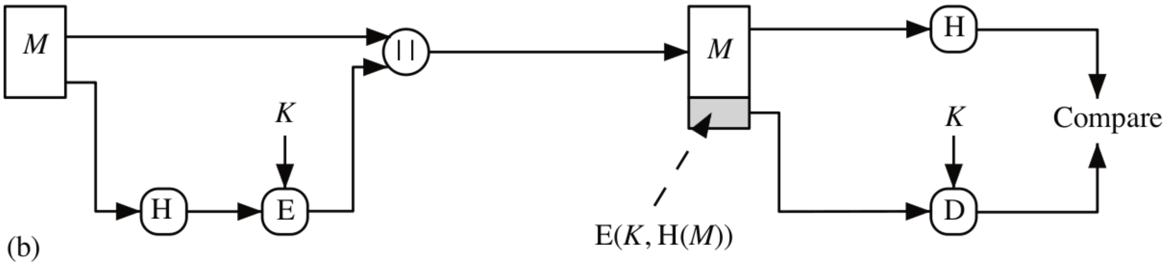

Figure 5: Security mechanism 2

(d) What is a security service that this mechanism provides? [1 mark]

Answer. *Authentication, data integrity*

(e) Explain (or define) the *weak collision resistant property* (also called *second preimage resistant property*) of a hash function. [2 marks]

Answer. *Computationally hard to find a message y such that* $H(x) = H(y)$*, given the hash function and x*

(f) Explain how an attacker can defeat the above security service if the function H() did not have the weak collision resistant property. [2 marks]

Answer. *If the weak collision resistant property does not hold, then the attacker can replace the message M sent by A with another message y, where* $H(M) = H(y)$*, and forward the message to B. B will not detect the change because after decrypting, the received hash value* $H(M)$ *will match the calculated hash value* $H(y)$*.*

(g) What is the difference between a hash function and a MAC function? [1 mark]

Answer. *A hash function takes data as input, while a MAC function takes data and a key as input*

(h) Explain what HMAC does when used with MD5. [1 mark]

Answer. *HMAC turns a hash function, MD5, into a MAC function*

Question 9 [18 marks]

You are developing a shopping website for a company. The website allows users to register (they are given a random, 6-digit user ID and can select any password between 8 and 10 characters in length, inclusive), login to obtained personalised content and services, as well as to purchase products and services using supplied credit card information. The company runs the web server, as well as a database server for storing user and product information.

(a) What protocol(s) should be used so that information transferred between users and the web server is confidential? [1 mark]

Answer. *HTTPS (or HTTP and SSL), as it provides encryption of data between web browser and server*

(b) The company has obtained a digital certificate issued by the authority VeriSign. Explain how this certificate can be used for web server authentication. (Include any assumptions about the web server or browser). [2 marks]

Answer. *The certificate is sent to the web browser. The web browser must have the certificate of the authority VeriSign. The browser uses VeriSigns certificate to verify the servers certificate, proving that the client is communicating with the intended server.*

(c) Certificates are generally not used for client (user) authentication. Explain then how client authentication is performed (including any assumptions). [2 marks]

Answer. *Once a secure connection is established between client and server, the user provides a username and password. The client is authenticated if the supplied username/password match the one selected during registration.*

(d) When a new user registers with the website, explain what identifying information must be stored in the database. [2 marks]

Answer. *At least the username/ID and and a hash of the password*

(e) Describe two methods you would implement that could prevent or deter online password guessing. For each of the methods, also describe the disadvantage of the method. [3 marks]

Answer.

- *Limit the number of incorrect attempts, e.g. to 10. This will be inconvenient for a user that forgets their password as they will not be able guess it.*
- *Introduce a delay between incorrect attempts, e.g. 5 seconds before another attempt can be made. Inconvenient for the user that types their password incorrectly.*

- *Log all incorrect attempts, reporting them to the user once they log in. Can only be used to track users after attempts, not while in progress.*

Consider the storage of user login information in the database. First assume that all passwords are chosen from the set of uppercase and lowercase letters. You are using MD5 128-bit, but not using a *salt* in the database. A malicious user has gained access to your database. They also have access to a rainbow table covering all passwords.

(f) Explain what information is stored in the rainbow table. [2 marks]

Answer. *The rainbow table stores all possible passwords and their corresponding hash values.*

(g) Assuming the rainbow table allows for compressing the raw data by a factor of 1,000,000, then how big would the rainbow table be? [2 marks]

Answer. *Passwords are between 8 and 10 characters in length (inclusive). There are* 52^8 *possible 8-character passwords, each taking 8 Bytes to store, plus 16 Bytes for the MD5 hash. The total size is* $52^8(8+16)$ *Bytes. Similar for the 9-character passwords and 10-character passwords. The total is then:*

$$52^8(8+16) + 52^9(9+16) + 52^{10}(10+16) = 3.8 \times 10^{18} Bytes$$

or about 3.8 million TB. But with compression, the reduces to 3.8 TB.

Now assume you decide to use a 20-bit *salt* value when implementing the registration/login system.

(h) Explain how the salt is chosen and what is stored in the database for each user. [2 marks]

Answer. *The salt is chosen randomly for each user, and the username, salt and hash of the salt concatenated with password are stored in the database.*

(i) Explain the main security benefit of using a salt, and how it provides that benefit. [2 marks]

Answer. *Using the salt prevents attacks using rainbow tables. A malicious user would need a rainbow table for each possible salt value. With a 20-bit salt, there are* 2^{20} *rainbow tables needed, effectively increasing the time to generate tham by 1 million and increasing the storage space required by 1 million, making it impractical to use.*