Final Exam

Network Security Autumn 2017

8 February 2018

Surname, Given Names (*e.g.*, Turing, Alan Mathison): _____

Student Identification Number (*e.g.*, 15-123-456): _____

Student Signature: _____

Rules and guidelines:

- Place your identification card on your desk. An assistant will check your identity during the exam.
- Once the exam starts, make sure you have received **all** pages of the exam. The exam should have **17 pages total**, including a page for extra space. **Do not** separate the exam sheets.
- Do not forget to fill in your **name, student identification number and signature** on this page.
- You **must** answer questions using **black or blue ink**. Illegible answers may not get any credit.
- The use of notes, textbooks or other written materials is **not** allowed. You are allowed to use a **scientific calculator** during the exam. Any other device that provides communication or document storage capabilities is **not** allowed (this includes smart watches).
- You have **90 minutes** to complete this exam.
- As a general guideline, one point should correspond to one minute. Thus you should write answers that are **clear and concise**. Generally, you do not need to completely fill the space provided for solutions.
- You are **not** required to score all points to get the maximum grade.
- When answering questions, always **explain your reasoning**. If a question asks, for instance, whether A is more secure than B, a plain "yes" or "no" answer will not be awarded any points.
- For questions during the exam, **raise your hand** and an assistant will come to answer your question.
- If you need extra space to answer a question, use the page provided at the back of the exam.
- At the end of the exam, please **remain seated** while we collect the exams. You may hand in your exam before the end, except in the last 10 minutes of the exam. Please **hand in all exam sheets**: if any sheet is missing, the examination will be marked with grade 1.0 and counts as failed.

Question:	1	2	3	4	5	6	7	8	Total
Points:	9	10	7	16	13	15	11	9	90
Score:									

1. TLS, PKIs, Certificate Transparency (9 points)

(a) (3 points) An attacker is trying to attack the company Wahoo and its users. Assume that users always visit Wahoo's website with an HTTPS connection, using fixed Diffie-Hellman and AES encryption. (You may assume that Wahoo does not use certificate pinning.)

For each of the following attack scenarios, mark *all* of the options that an attacker could achieve in that attack scenario. (Partially correct answers and incorrect answers get 0 points.)

i. (1 point) If the attacker obtains a copy of Wahoo's certificate, the attacker could:

- ☐ impersonate the Wahoo web server to a user
- ☐ discover some of the plaintext of data sent during a past connection between a user and Wahoo's website
- ☐ discover all of the plaintext of data sent during a past connection between a user and Wahoo's website
- ☐ successfully replay data that a user previously sent to the Wahoo server over a prior HTTPS connection
- √ **none of the above**

ii. (1 point) If the attacker obtains the private key of a certificate authority trusted by users of Wahoo, the attacker could:

- √ **impersonate the Wahoo web server to a user**
- ☐ discover some of the plaintext of data sent during a past connection between a user and Wahoo's website
- ☐ discover all of the plaintext of data sent during a past connection between a user and Wahoo's website
- ☐ successfully replay data that a user previously sent to the Wahoo server over a prior HTTPS connection
- ☐ none of the above

iii. (1 point) Suppose the attacker obtains the private key that was used by Wahoo's server during a past connection between a user and Wahoo's server, but not the current private key. Also, assume that the user has learned that the certificate corresponding to the old private key has been revoked and is no longer valid. This attacker could:

- ☐ impersonate the Wahoo web server to the user
- ☐ discover all of the plaintext of data sent during a current connection (one where the current private key is used) between the user and Wahoo's website
- √ **discover all of the plaintext of data sent during a past connection (one where the old private key was used) between the user and Wahoo's website**
- ☐ none of the above

(b) (2 points) For server efficiency reasons, many banks prefer not to set up TLS connections for the page containing username/password entry, but instead create a non-TLS page that contains Javascript code that encrypts the username/password before sending it to the server. Provide a brief argument on whether this approach is secure or not.

Solution: Not secure because the adversary can modify the Javascript code and perform MiTM attack to retrieve the username/password.

Grading Guideline: -2pt if say yes, -1.5 pt if incorrect reasoning, -1 pt if the answer does not explicitly mention compromise in username/password (e.g., just mention MiTM attack)

(c) (4 points) Recently, a study has shown that many governments are forcing Certificate Authorities to issue them forged certificates which will enable them to carry out man-in-the-middle attacks against TLS-secured web sites. Consider the following hypothetical example of such an attack:

The CIA wants to monitor the activities carried out by the customers of Commercial Bank of Dubai (whose servers reside in Dubai). The CIA forces VeriSign to generate a valid certificate for the name "Commercial Bank of Dubai" (whose actual certificate is issued by Etisalat, UAE). Now, the CIA can perform a man-in-the-middle attack against all the US users who access the Commercial Bank of Dubai from within US. This attack is undetectable by current browsers as VeriSign is already a trusted CA.

This example shows what is known as the *compelled certificate creation attack*. Answer the following questions about this attack:

i. (2 points) What is the risk taken by the government (or by the CA that issues the second valid certificate) when they try to carry out such an attack against users?

Solution:
Since the certificates are non-repudiable, anyone who gets hold of the second valid certificate signed by VeriSign can prove VeriSign's (or government's) involvement in the attack.
However, if the CA creates an intermediary CA certificate for the government, then that certificate would have to be registered in a CT log, but in that case, the attack does not become directly attributable to VeriSign, as the forged certificate is signed by the intermediary certificate.
Grading Guideline: -1pt if does not mention loss of trust of the govt or the CA. -2pt for incorrect or other reasonings.

ii. (2 points) Can Certificate Transparency be used to make such an attack publicly detectable? If yes, under what circumstances? If no, why not?

Solution:
Yes, if the forged certificate is registered with CT, then the attack becomes trivially publicly detectable. If the forged certificate is not registered, then the browser would have to report the certificate, which would make the attack public.
Grading Guideline: -1 pt if correct justification but does not explicitly state that the two certificates must be logged. -1.5 pt for correct answer but incorrect justification. -2 pt for incorrect answer.

2. DoS and Botnets (10 points)

Consider an attacker that wants to perform a DDoS attack on a victim server. The attacker has created a botnet with 10,000,000 compromised IoT devices but is *unable to spoof source addresses*.

(a) (1 point) The attacker uses the botnet devices to directly flood the victim server. If each IoT bot device simply sends 10 TCP SYN packets per second towards the victim, what is the approximate aggregate bandwidth that reaches the victim? Give your answer in the unit of bits per second.

Solution: TCP has a minimum header size of 20 bytes, IPv4 a minimum of 20 bytes and IPv6 a minimum of 40 bytes. Thus we have $10^1 \times 10^7 \times 8 \times s = 32 \times 10^9$ bps (48×10^9 bps) minimum for IPv4 (v6) + TCP headers. **1 point** for accurate assumptions or calculations assuming only addresses and ports. **0.5 points** for plausible size assumptions outside of the range, unless otherwise justified.

(b) (3 points) Consider a scenario where the victim server uses an address filtering approach. Botnet nodes are blacklisted (based on their regular connection requests), and the associated packets are rejected. What is an advantage of this approach and a limitation?

Solution: This would prevent the server having to process any of the malicious SYN requests **(1.5 points)**. If the congestion is in the network, then the regular packets cannot reach the server; blacklisting a NATed IoT device can cause collateral damage **(1.5 points)**. **0.5 points** for implementation disadvantages such as performance/complexity arguments.

(c) (3 points) Consider a network device that can store up to 10,000 address filtering rules. How could this device be used to mitigate the attack described in part (a)? Describe a disadvantage of the suggested approach?

Solution: Addresses can be aggregated into and filtered using subnets **(1.5 points)**. This will cause false positives (collateral damage) for legitimate hosts who are unlucky to end up in a traffic aggregate, as well as false negatives for botnet hosts that do not end up in a traffic aggregate **(1.5 points)**.

(d) (3 points) Now assume that botnets are *capable of spoofing addresses*. Consider a service which publishes a list of misbehaving hosts, categorized as those performing 20 consecutive SYN requests to one of the servers in the network controlled by the service. Any server may periodically download the blacklist and refuse connections by listed hosts. Describe two disadvantages of this approach.

Solution: 1.5 points for each motivated disadvantage. Disadvantages include (i) the service can be DDoSed, (ii) an attacker knowing the scheme can avoid being placed on the list, (iii) change of IP address after being blacklisted, (iv) framing of legitimate users by spoofing their addresses.

3. Intrusion Detection, Firewalls and Evasion (7 points)

(a) (2 points) Next generation firewalls utilize application and protocol semantics to filter malicious traffic. Describe a possible negative impact of such devices being pervasively deployed in the Internet.

> **Solution: 2 points** for any of the following: (i) this can lead to ossification of the protocol to port assignments, (ii) restrict the development of new protocols, (iii) no benefit for encrypted traffic. **1.5 points** for performance arguments. **1 point** for privacy arguments, as relevant entities already have access to the network traffic.

(b) (3 points) Consider an IDS that uses a hash table to store flow identifiers. The hash table is implemented with an MD5 cryptographic hash function applied on the flow's 5-tuple which returns an array index, for example (where % represents the modulo operator):

```
index = md5(src_ip, src_port, dest_ip, dest_port, protocol) % array_size
```

Each entry of the array contains a linked list of pointers to handle colliding entries.

i. (2 points) How could an attacker exploit this data structure to perform a DoS attack on the IDS?

> **Solution:** An adversary could carefully select the parameters for the md5 function to create colliding entries **(1 point)**. This results in a long linked list in this slot, thus slow performance in traversing the list **(1 point)**.

ii. (1 point) How can this attack be mitigated?

> **Solution:** Using a cryptographic hash function with a random secret key **(1 point)**. **0.5 points** for use of randomness. The collision space is the array size, to make the array large enough to prevent frequent collisions would be infeasible, changing the hash function is insufficient **(0 points)**.

(c) (2 points) Your company currently deploys an IPS system on the office network. Some of employees of your company have requested to be allowed to connect to their home networks via VPN from the office. Why is this not safe to allow? Describe a mitigation for the perceived risk.

> **Solution:** Threats may enter the office network through the VPN. As the traffic is encrypted, the IPS cannot analyze it **(1 point)**. A host based IPS could be used, or the IPS could be provided with the keys for decrypting the VPN **(1 point)**.

4. Anonymous Communication Systems (16 points)

(a) (4 points) **Tor: compromised directory.** Directory authorities are very powerful entities in Tor, since they regularly publish the list of all Tor relays and their status. The authorities work by a consensus algorithm such that a majority of directory authorities can change the list that is being published. Assume that an adversary is able to compromise 6 out of the 10 directory authorities in Tor:

i. (2 points) Describe one way in which this adversary can de-anonymize a large fraction of the Tor network.

> **Solution:** The adversary may change the flags and weights of the relays in the consensus document in a way that the relays under the adversary's control are highly likely to be chosen when creating circuits **(2 points)**.
> Alternatives: the adversary can create different consensus documents to give to the caching relays, creating different views of the networks for users using different entry guards **(2 points)**.
> Significantly less effective attacks, or attacks that are not clearly explained, get **1 point**.

ii. (2 points) Which entity (or entities) may detect the attack you described? (Motivate your answer.)

> **Solution:** The other 4 directory authorities can see that consensus documents are being signed which are significantly different/inconsistent with the reports they have received from the relays. Also other relays and users may notice that the consensus document they receive changes significantly/rapidly, which could raise alarms. Furthermore, relays may notice that they're not listed anymore.
> An answer with a single entity gets **1.5 points**, with two entities or more **2 points**. Deduct **0.5–1 points** for insufficient motivation of the answer, or for a detection mechanism which would be ineffective/impractical.

(b) (4 points) **Tor: compromised website.** With reference to the previous question, consider now a different adversary which is able to obtain the private key for the certificate of `torproject.org`, the website used to distribute the Tor client (browser) and all the updates to it. Describe (at least) one advantage and one disadvantage of this adversary's capabilities compared to the previous adversary which could compromise 6 out of 10 directory authorities. Which adversary is stronger, and why?

> **Solution:** The ability to compromise the distribution of the Tor client and its update process implies the ability to fully compromise any user that downloads and uses a malicious client: the malicious client can simply forward all metadata about the visited websites to the adversary **(1.5 points)**. Being able to determine the client's code also means that the adversary can change the list of directory authorities, e.g., to include only server's under the adversary's control. In this sense, this adversary can be considered strictly more powerful than the previous one **(1 point)**. (Implicit answers to this may get 0.5 points.)
> A disadvantage of this adversary's capabilities is that an attack based on modifying the code of the Tor browser can be detected more directly by security-conscious users which check the hash of the binary before updating by comparing it with those distributed by Tor developers and users on Twitter and other channels **(1.5 points)**.
> Alternatives: Another advantage is that the adversary can see the plaintext even when HTTPS is used; another disadvantage is that this adversary would need MitM capabilities to mount the attack. If well argued, an answer saying that the first attacker is stronger (e.g., because no MitM is needed) can be accepted.

(c) (8 points) **Tor: long circuits.** In Tor, a `create` cell is sent to a Tor relay to establish new keys and make that relay be part of a circuit that is being created. For instance, a `create` cell is sent by a client to an entry guard to establish the first hop of a new circuit. To add a hop to a circuit that is being created, a client sends a `relay_early` cell, which carries an `extend` command to the (current) last hop. This last relay then uses the information in the `extend` command to send a `create` cell to the relay that should be added to the circuit.

The use of `relay_early` cells instead of the normal `relay` cells (used to transport normal data) has been added to Tor to prevent the creation of very long circuits. Relays only accept `extend` commands when they are contained in `relay_early` cells, and any relay will accept only up to 8 `relay_early` cells for a single circuit.

For the questions below, we define the *amplification factor* of a circuit as the number of (data) cell processing actions performed by honest relays divided by the number of (data) cells sent by a malicious sender. (For example, a default circuit consisting of three honest relays has an amplification factor of 3.)

i. (2 points) What is the maximum amplification factor for a simple circuit? (Without using any tricks; assuming all Tor relays are honest.) Briefly explain your reasoning.

> **Solution:** The maximum amplification factor is 9, because the entry guard is the relay seeing the most `relay_early` cells, and it will allow only 8, meaning that, *apart from the entry guard itself*, up to 8 relays can be added to the circuit **(2 points)**. Deduct **1 point** for insufficient explanation.

ii. (3 points) Do you see ways to increase this factor? If so, by how much? (Assume that all Tor relays are honest.)

> **Solution:** One way is by using hidden services, which will increase the number of honest relays traversed by a single cell **(1 point)**. When connecting to an honest hidden service, the amplification factor increases to 11–13, when connecting to a hidden service controlled by the destination it can increase to $\geq$16 **(2 points)**. (Only one of the two options gets 1.5 points.)
> Alternative: it may be possible to tunnel Tor over Tor, i.e., to connect an exit node to a guard, and establish a new circuit over the existing circuit **(2 points)**. In theory this can be used to get an arbitrarily large amplification factor **(1 point)**.
> Providing at least two ways for increasing the factor compensates for not providing a (correct) value of the amplification factor.

iii. (3 points) Mallory thinks she has found a great way to circumvent the circuit length limitation mechanisms and perform powerful DoS attacks against Tor. She sets up a malicious relay, and builds a circuit that has her malicious relay as the 8th hop. Now on the first few hops she only uses `relay` cells, even when they contain `extend` commands, and she instructs her malicious relay to change the type of such cells from `relay` to `relay_early`. She believes this can lead to arbitrary circuit length if she places her malicious relay also as the 16th, 24th, ..., hops on the circuit. Is she correct? Would this lead to an arbitrarily high amplification factor? Explain your reasoning.

> **Solution:** While it is true that in this way a circuit of arbitrary length could be created **(2 points)**, it does not provide any additional amplification besides the one obtained by a simple circuit (7–9), since the malicious relay could just as well be establishing a new circuit for every time Mallory's circuit passes through the relay **(2 points)**.
> Deduct **-0.5 points** if it is not explicitly stated that she is correct, but it can be inferred from context.

5. Probabilistic Traffic Monitoring (13 points)

(a) (7 points) A *wireless ISP* charges the users (with individual IP addresses) of its *WiFi* service offering based on their bandwidth usage. As the number of users is large, the company decides to forgo precisely billing every user to enable the use of commodity hardware. Using Netflow, they instead sample every k-th packet, and bill the sampled customers according to the inferred usage. While aware that some users may luckily avoid a charge, they anticipate that users which use high bandwidth or the network for prolonged period will likely be charged.

i. (1 point) Give a definition of "flow" useful in the billing process.

> **Solution:** We have both incoming and outgoing data. On the incoming data a flow would be a set of packets sharing the same destination IP address **(0.5 points)** and on the outgoing those sharing the source IP address **(0.5 points)**. **0.5 points** for excessive but accurate definitions of flows.

ii. (2 points) Can this billing approach be considered fair for paying users with respect to their usage? Why or why not?

> **Solution:** No, it's possible to over-estimate the usage **(1 point)**. The estimate would be $k * b$, for b bytes sampled. **1 point** for any explaining how it is oversampled: (i) an unlucky user whose sole packet was sampled would be over-charged, (ii) the sampling method is biased towards more frequent packets versus the actual size of the packets.

iii. (2 points) Describe an attack that could be launched against such a system.

> **Solution: 2 points** for any of the following attacks: (i) since the system is over WiFi, an adversary knowing the sampling rate could avoid being sampled, (ii) the adversary could fill the interval so as to increase the likelihood of another user being sampled, especially for users transmitting bursts of packets.

iv. (2 points) The company instead decides to focus on billing the users consuming the most bandwidth. Suggest an approach for collecting flow throughput.

> **Solution:** Sample-and-hold would make best use of the available space. Large flows are sampled with a higher frequency and no over-estimations are made. Some small flows will also be sampled and can be billed **(2 points)**.
> Frequent item finding and majority algorithms give worse approximations of the bandwidth as the tracked values are decremented, but could also be used **(1 point)**. Multistage filter could also work, but has false positives/over-estimations **(0.5 points)**.

(b) (6 points) You are responsible for the network of a company which serves two resources, X and Y, from two different web servers. Resource X is 2 MB whereas resource Y is 5 KB. You have asked for the outgoing link of 100 Gbps to be upgraded, as flows serving resource X dominate the outgoing link resulting in an increased latency for flows serving resource Y. Unfortunately, you do not have access to the application logs to prove this to your supervisor.

i. (2 points) Describe an approach for measuring the consumed throughput of each resource, assuming the router does not have sufficient resources to individually track each flow.

Solution: Since the resources reside on different servers, we can track flows by the source address of the server to measure the bytes sent within some interval t **(2 points)**. **1 point** for suboptimal/wasteful/approximate but correct solutions.

ii. (4 points) Consider the case when both resources instead reside on the same server. How can you determine the exact number of flows for resource X and at least an approximate of the number of flows for resource Y within an interval T?

Solution: Since the two resources reside on the same server, we must use the sizes to distinguish them **(0.5 points)**. We can use EARDet-like algorithm (without virtual traffic) and define the ambiguity region to be the interval between the two resource sizes. This will ensure that none of resource Y are counted as resource X and as there are no resources whose size are in the ambiguity region, there would be no false positives **(2.5 points)**. We can use probabilistic counting to determine the total number of unique flows, and take the difference for an estimate of the number of small flows **(1 point)** **1 point** for any probabilistic algorithm that can be used to estimate the number small flows.**1 point** for any inexact estimates of X.

6. DNS and DNSSEC (15 points)

(a) (4 points) **Caching resolvers.** Recursive resolvers have an important role in DNS: by caching the replies they obtain, they significantly lower the average lookup latency for end hosts using the resolvers. However, the use of caching is not without its drawbacks.

i. (2 points) Name and describe one drawback in terms of *privacy*.

> **Solution:** By measuring the latency of replies from a resolver, an adversary can figure out whether certain names have been queried by other clients using that resolver **(2 points)**.
> Alternative: the resolver itself is an additional entity learning what queries the user is making **(1.5 points)**, which is a problem in particular for resolvers that are not provided by an ISP **(0.5 points)**.

ii. (2 points) Why would an adversary prefer a world in which caching resolvers are used extensively (as they are in our world) over a world in which all end hosts' clients perform the full recursive look-ups on their own?

> **Solution:** Because it allows cache poisoning attacks **(1.5 points)** that have potentially a large impact, as many clients would be affected **(0.5 points)**.
> Alternatives: Because it would put a heavier load on the DNS infrastructure, making DoS attacks potentially easier **(2 points)**. Because, if the resolvers are configured to be open **(0.5 points)**, they can be leveraged for amplification attacks **(1.5 points)**.

(b) (4 points) **Resolvers and DNSSEC.** Recall that DNSSEC is typically used only between recursive resolvers and authoritative name servers. Assuming DNSSEC is fully deployed, consider an end host that fully trusts the recursive resolvers provided by its ISP (so it also trusts that the resolver will always use DNSSEC).

i. (2 points) What attack undermining the authenticity of the query replies should the end host still worry about, and why?

> **Solution:** The link between the end host and the resolver is in general not secure, so an eavesdropping attacker (e.g., in the same LAN network) may alter the replies **(2 points)**.
> Alternative: an adversary could even replace the resolver entirely by hijacking the initial DHCP setup **(2 points)**.

ii. (2 points) What countermeasure (if any) could be adopted to prevent or significantly mitigate the attack you mentioned? Briefly explain your answer.

> **Solution:** The end host could set up an authenticated channel to the resolver, ensuring (via certificates or other) that it is getting the replies sent by the correct resolver **(2 points)**.
> Alternative: the end host could do the entire DNSSEC lookup on its own **(2 points)**.

(c) (7 points) **Authenticated denial of existence.** DNSSEC is not only concerned with the authentication of the usual DNS records; it also aims to authenticate the *absence* of certain records. At a high level, the strategy DNSSEC adopts to do this is the following: to authenticate the non-existence of, e.g., `medicine.ethz.ch`, the authoritative server needs to provide a signed non-existence statement containing the pair of existing names which alphabetically come respectively right before and right after the non-existing name. For instance, this pair could be (`mavt.ethz.ch`, `mtec.ethz.ch`).

Concretely, DNSSEC introduces the NSEC record, which could look as follows:

NAME	TTL	CLASS	TYPE	NEXT DOMAIN NAME	TYPE BIT MAP
mavt.ethz.ch	3600	IN	NSEC	mtec.ethz.ch	(A, MX, RRSIG, TXT)

i. (2 points) The *Type Bit Map* shown above specifies what record types exist for `mavt.ethz.ch`: why is this field needed/useful in the context of authenticated denial of existence?

> **Solution:** It allows the authenticated denial of existence certain types of records for names for which records of other types exist **(2 points)**.

ii. (2 points) What other records from the `ethz.ch` domain, besides the NSEC record shown above, are needed to verify the non-existence of `medicine.ethz.ch`?

> **Solution:** The `DNSKEY`s of the `ethz.ch` zone (KSK and ZSK) are needed **(1 point)**, and the signature (`RRSIG`) over the NSEC record **(1 point)**.

iii. (3 points) A much more trivial solution would be for the authoritative name servers to sign non-existence statements on the fly for each query they receive for non-existent records. Name *two* significant drawbacks of this solution, and briefly explain your answer.

> **Solution:** Adopting this solution would mean that authoritative name servers have to perform online asymmetric crypto operations, which is very costly, and would make those servers highly vulnerable to DoS attacks **(1.5 points)**. Furthermore, this would also require the name servers to keep the signing keys online, which makes them more vulnerable to compromise **(1.5 points)**.
> Alternative: it would increase the latency of replies, as the name servers have to compute the signatures, which takes longer **(1.5 points)**.
> Answers that focus on issues *which affect NSEC as well* get **0–0.5 points**. This is because these issues are mostly intrinsic "limitations" of the solution space of authenticated denial of existence (to the best of our knowledge), so they cannot really be considered "drawbacks" of the approach described in the question.

7. Broadcast authentication (11 points)

(a) (1 point) Is it possible to perform authenticated broadcast communication, assuming that the set of receivers may change over time, that the receivers have no prior information about the sender, and that any trusted third party is offline (i.e., a PKI CA)?

Solution:
We accept any variants of the following two answers. 1) No, need a root of trust for authentication. 2) Yes, include certificate that was issued by the trusted third-party in the broadcasted message.
Grading Guideline: Binary. No partial point awarded.

(b) (1 point) Alice wants to send messages to mutually untrusted receivers Bob, Carol, and Dave. Alice picks a random key K and sends it to the receivers over a secure channel providing secrecy and authenticity. To broadcast message m, Alice sends out m, MAC(K, m). Since each receiver knows key K, each receiver can verify the authenticity of the message by verifying the MAC. Assuming the MAC function is secure, is this protocol secure?

Solution:
No, because Bob/Car/Dave can impersonate Alice.
Grading Guideline: Binary. No partial point awarded.

(c) (9 points) Answer the following questions about the TESLA protocol.

i. (2 points) The sender has been broadcasting with TESLA to a large group, and it turns out that the sender used all keys of the original hash chain. Is it possible to extend the one-way hash chain and continue the TESLA broadcast operation efficiently? Justify your answer.

Solution:
No. It's necessary to create a new hash chain, and the hash chain is used in the reverse order of construction so you cannot extend it.
Grading Guideline: -1pt If answering yes with correct reasoning about the extension; however, if drawback of such extension not mentioned. -1.5pt If answered No, but with incorrect justification. -2pt for incorrect answer.

ii. (4 points) An important TESLA parameter is the key disclosure delay. Although the choice of the disclosure delay does not affect the security of the system, it is an important performance factor. As we discussed in class, a short disclosure delay will cause delayed packets to lose their safety property, so receivers will discard them, and a long disclosure delay leads to a long authentication delay for receivers.
As an alternative, the sender may include in each packet the time t_p at which it is going to disclose the key for this packet. With this method, the receiver only needs to know the bound D_t on the clock skew and T_0, the sender's local time at the initiation of the session. Then the receiver records the local time T when the packet has arrived, and verifies that $T \leq T_0 + D_t + t_p$. Else the packet is considered unauthenticated. Is this secure? Justify your answer.

Solution:
An attacker can capture the correct packets, wait for the key disclosure, and re-send forged packets with delayed key disclosure times. The scheme is thus trivially broken.
Grading Guideline: -3pt if just mentions DoS Attack or provides incorrect justification. -2pt for vague reasoning but on right track, -1pt for almost correct but does not clearly state the attack. -4pt for incorrect answer.

iii. (3 points) Instead of operating on a time basis, sender S decides to operate TESLA on a packet basis. S now broadcasts the packet P_i along with the key K_i and the message authentication code (MAC) of P_i computed with the key K_{i+1} as follows:

$$S \rightarrow * : P_i, K_i, MAC_{K_{i+1}}(P_i)$$

The receiver must wait for the next packet to validate the MAC of the packet P_i.
Is this secure? Justify your answer.

Solution:
No, this is not secure. Since keys are not released on a time basis, an attacker can forge messages. For instance, an attacker may grab two subsequent packets (P_i and P_{i+1}) from S, replace P_i with its own message $P_{i'}$ and use K_{i+1} to compute MAC. When the attacker rebroadcasts $P_{i'}$ and P_{i+1}, receivers cannot tell whether $P_{i'}$ is forged or not.
Grading Guideline: -3pt for incorrect answer, -2pt if answered non-secure but wrong or vague description, -1pt if does not completely describe the attack.

8. SCION (9 points)

In this question, we will explore the possibility of various attacks against SCION. NOTE: For those who do not remember the details of the SCION protocol from the lecture slides, we have reproduced the relevant data structures on the next page.

(a) (2 points) **Hop field spoofing attack**. Consider a malicious host that desires to send packets down a link for which it does not have a valid hop field. How can it create a packet that traverses that link?

> **Solution:** The attacker can brute force the MAC, until a packet gets through. Since the MAC is 24 bits long, it would require around 8 million packets on average.
> **Grading Guideline: -2pt if answer no, -1pt if wrong or vague reasoning**

(b) (3 points) **Reflection attack.** In this question, we explore if a reflection attack can be launched in SCION.

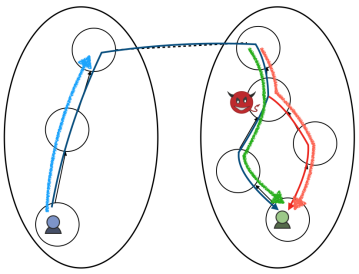
i. (1 point) Can an adversary launch a reflection attack against a victim that is in the same AS? Justify your answer.

> **Solution:** Yes, just use the IP address of the victim as the source address. OR No, use DRKey or ingress/egress filtering.
> **Grading Guideline: Binary. Based on the reasoning.**

ii. (2 points) Can an adversary launch a reflection attack against a victim in another AS? Justify your answer.

> **Solution:** Yes (in some cases). If the adversary is on the path between the victim and the reflection point, then the adversary can launch a reflection attack, by setting the hop field pointer to the current AS and by sending the packet to the egress border router.
> OR No, use DRKey.
> **Grading guideline: -2pt if just answered yes or no. OR w/o or incorrect justification. -1pt if it provides a vague reasoning.**

(c) (2 points) **Path alteration attack** occurs when a malicious entity, such as a compromised border router in an AS, modifies the SCION path in a SCION data packet, and the packet reaches the intended destination via the modified path. Is SCION resilient against a path alteration attack? Justify your answer.

> **Solution:** No, in the figure below, the adversarial AS can alter the path (in blue) between the source (in blue) and destination (in green) hosts. Specifically, the adversary can replace the hop fields for the last three ASes with the hop fields from the red down-segment so that the packet would traverse through the red path to the destination host (in green). Since a packet would arrive at the destination using a path that is different from what the source host has specified, the attack is successful.
>
> 

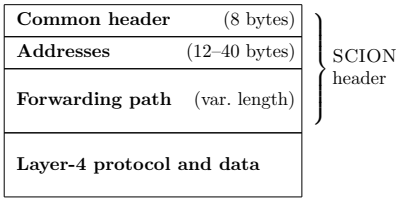
Grading guideline: -2pt if wrong answer. -1pt if wrong or vague justification.

(d) (2 points) **Wormhole attack** occurs when two ASes collude to announce a bogus link between them. This consequently creates shorter paths to attract traffic. Does SCION defend against a wormhole attack? Why or why not? Justify your answer.

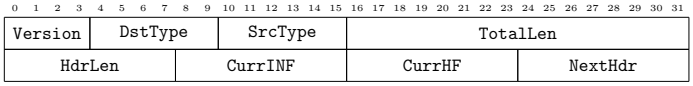
Solution: No. Add a fake link to the beacon messages.
Grading guideline: -2pt if wrong answer. -1pt if wrong or vague justification.

Appendix: SCION Information

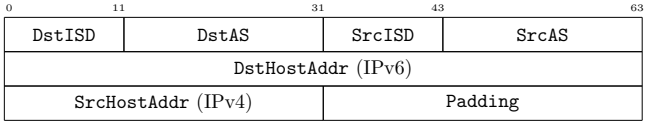
A **SCION data packet** has the header structure as shown below, and we provide more details for the three fields (i.e, **SCION Common Header, Addresses, Forwarding paths**) that constitute the **SCION header**.



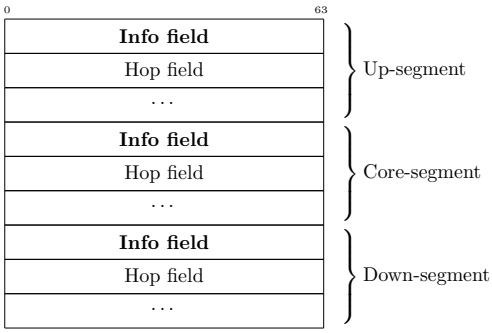
The **SCION common header** has the following structure:



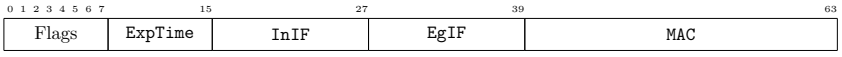
The **Addresses** field indicates the addresses of the source and destination hosts. Assuming that the source and destination ASes are an IPv6 and an IPv4 networks respectively, the addresses field has the structure as shown below. Note that a complete SCION address is defined by a (ISD Number, AS Number, and Local Address)-triplet.



The **Forwarding Path** field contains information to forward packets across ASes and has the following high-level layout:



A **Hop Field** in a forwarding path has the following structure:



The only cryptographic aspects in a SCION header are the MACs in the hop fields in a forwarding path. Each MAC, which is 24 bits, is created by each AS during the beaconing process and is computed as follows:

$$\sigma_H = \mathsf{MAC}_K(TS \parallel Flags \parallel ExpTime \parallel InIF \parallel EgIF \parallel HF')$$

In this equation, TS refers to the timestamp (in the Info field), Flags to the hop field flags, ExpTime to the expiration time, InIF and EgIF to the ingress and egress interfaces, and HF' is the hop field of the AS from which the beacon was received (it is empty when the core AS creates the first hop field).

Extra Page

Please use this page in case you run out of space elsewhere in the exam.