CS 161 Summer 2024 | Introduction to Computer Security | Exam Prep 13

Q1 *Networking: A TORrible Mistake* (7 points)

Q1.1 (1 point) Assuming no malicious nodes collude, an n-node Tor circuit provides anonymity (i.e. no node learns who both the user and server are) when at least _____ node(s) are honest. Fill in the blank.

- ● 0
- ● 1
- ○ $n - 1$
- ○ n

Solution: The intended answer was 1. As seen in lecture, a Tor circuit is secure if at least one node is honest. Anonymity is only broken if every node in the circuit colludes, so that together they can reconstruct the entire circuit that messages are being routed through.

However, after the exam, we decided the question wording was unclear, because it assumes that no malicious nodes collude. If no malicious nodes collude, then Tor is secure, even if none of the nodes are honest, so we accepted 0 as an alternate answer.

For the next 3 subparts, a user is using Tor to send a message to a server. Assume that there is no collusion between any Tor nodes, and that the user choses exactly 3 nodes for their Tor circuit.

Q1.2 (1 point) Which values can a malicious **entry** node learn? Select all that apply.

- ■ The IP address of the user
- □ The IP address of the server
- □ The list of all nodes in the circuit
- □ None of the above

Solution: The user sends messages to the entry node, telling the entry node to forward those messages to the next node.

The IP address of the server is wrapped in many layers of encryption inside the message sent to the entry node, so the entry node cannot see that value.

The entry node knows about the second node in the circuit, but not the entire list of nodes.

Q1.3 (1 point) Which values can a malicious **exit** node learn? Select all that apply.

☐ The IP address of the user

■ The IP address of the server

☐ The list of all nodes in the circuit

☐ None of the above

> **Solution:** The exit node is the last node in the circuit, who needs to know the server's identity so that they can forward the message to the server.
>
> By the time the message reaches the exit node, all information about the original user's identity has been stripped away (the entry node removed all traces of the original user's identity when forwarding the packet to the second node).
>
> The exit node knows about the second-to-last node in the circuit, but not the entire list of nodes.

Q1.4 (1 point) Which values can an on-path attacker on the user's local network learn? Select all that apply.

■ The IP address of the user

☐ The IP address of the server

☐ The list of all nodes in the circuit

☐ None of the above

> **Solution:** The on-path attacker in the local network can see the user sending messages into the Tor network (to the entry node).
>
> However, the IP address of the server is encrypted inside the message sent to the entry node, so the on-path attacker cannot see that value.
>
> The on-path attacker only knows about the entry node, not the entire list of nodes in the circuit.

When a new user first downloads Tor, they need to download a list of nodes from a trusted directory server.

A malicious, on-path attacker on the user's local network wishes to eavesdrop on the new user's Tor connection. Assume that the attacker controls 3 nodes out of 100 total Tor nodes, and can win any data race.

For the next three subparts, select the approximate probability that the attacker can learn the identity of the server.

Q1.5 (1 point) User connects to the directory via TLS, attacker is on-path.

- ○ Exactly 0%
- ● Greater than 0%, less than 50%
- ○ Greater than 50%, less than 100%
- ○ Exactly 100%

Solution: Because the directory connection is made over TLS, and TLS has end-to-end security, the on-path attacker cannot tamper with the list of nodes.

Therefore, the on-path attacker can only hope that the user randomly selects the three nodes controlled by the attacker.

The probability of selecting the 3 attacker-controlled nodes out of 100 nodes is intuitively less than 50%, but it's not 0%.

Formally, you can calculate this probability to be $6/(100 \cdot 99 \cdot 98)$, where the numerator is the number of ordered ways to choose the 3 attacker nodes (counting all possible orders, since order doesn't matter), and the denominator is the number of ordered ways to choose any 3 nodes.

Q1.6 (1 point) User connects to the directory via TCP, attacker is on-path.

- ○ Exactly 0%
- ○ Greater than 0%, less than 50%
- ○ Greater than 50%, less than 100%
- ● Exactly 100%

Solution: Unlike the last subpart, the user is now using just TCP to connect to the directory, so the attacker can tamper with the response from the directory.

Specifically, the attacker can trick the user into thinking that the list of nodes only has 3 nodes: the attacker-controlled nodes.

Now, the user is forced to always choose the attacker-controlled nodes, and the attacker will always be able to break anonymity by controlling every node in the resulting circuit.

Note that we don't have to worry about data races, since the question says the attacker can win any data race.

Q1.7 (1 point) User connects to the directory via TCP, attacker is off-path.

- ○ Exactly 0%
- ● Greater than 0%, less than 50%
- ○ Greater than 50%, less than 100%
- ○ Exactly 100%

Solution: As in the previous subpart, the attacker can trick the user into using the attacker's nodes.

However, because the attacker is now off-path, they need to guess the sequence number in order to inject a malicious message into the TCP connection. The probability of the attacker guessing a valid 32-bit sequence number is under 50% (but not 0%).

Q2 *Suit of Armor Around the World (SP22 Final Q8)* (16 points)

You are tasked with securing The Avengers' internal network against potentially malicious protocols! For each type of firewall and set of traffic, state whether the firewall is able to achieve the desired functionality with perfect accuracy. **Assume that IP packets are never fragmented.** All connections that are not mentioned can be either allowed or denied.

If you answer Possible, briefly (in 3 sentences or less) how the firewall should operate to achieve the desired effect. If you answer False, provide a brief justification for why it isn't possible.

Q2.1 (4 points) **Desired Functionality:** Block all inbound TCP connections. Allow all outbound TCP connections.

Firewall: Stateless packet filter

● Possible ○ Not possible

> **Solution:** This is possible by blocking all inbound packets with only the SYN flag set, which prevents inbound connections. This allows outbound connections by allowing outbound SYN packets, and the resulting inbound SYN-ACK packet is allowed.

Q2.2 (4 points) **Desired Functionality:** Allow all outbound TLS connections. Block all outbound TCP connections that aren't running TLS.

Firewall: Stateful packet filter

○ Possible ● Not possible

> **Solution:** While a stateful packet filter *can* reassemble a TCP data stream and look for signatures of a TLS handshake, it can still be circumvented with techniques such as sending multiple small TCP segments with the same sequence number but differing TTLs.

Q2.3 (4 points) **Desired Functionality:** Allow outbound DNS requests. Block inbound DNS responses. Assume that name servers always listen on port 53.

Firewall: Stateless packet filter

● Possible ○ Not possible

> **Solution:** This is possible (although it doesn't achieve much). One would allow outbound UDP datagram packets with the destination port 53 but block inbound UDP datagram packets with source port 53.

Q2.4 (4 points) **Desired Functionality:** Block all HTTP traffic that contains the literal string **Ultron**. Allow all other HTTP traffic.

Firewall: TCP proxy

● Possible ○ Not possible

Solution: TCP proxies allow the TCP stream to be reconstructed exactly. Once the stream is reconstructed, the firewall can keep track of the entire HTTP request as state and, if it contains the string `Ultron`, drop the connection.