Advanced Web Security - Assignment 2A

November 2020

Questions

1. Motivate why the lecture notes define anonymity roughly as “the IP is unknown”.
2. In the DH key exchange, the entry node does not encrypt g^x when it sends its ”created cell” handle back to Alice. Why no encryption is needed? If one wishes to encrypt it, what key should be used?
3. Describe the notion of ”perfect forward secrecy” (PFS) and why is it important for systems where the key material evolve. How does PFS transform in case there is no progress in the key material?
4. In a 2-mix chain, can mix1 learn the identity/IP-address of both the sender and the receiver of a message it receives?
5. When using 2 mixes and an untraceable return address, show how the addressee prepares the return message to the original sender.
6. When Alice creates a Tor circuit, who selects the relays that are used?
7. When negotiating a symmetric key with an onion router, what is the purpose of the H(K1) message sent from the router to Alice?
8. A TCP handshake consists of the client and the server exchanging three messages: SYN, SYN-ACK and ACK. Explain why, in Tor, Alice can connect to a web-server and expect the TCP handshake with the server to be performed with low latency.

Answers

1. While browsing online or using the internet in general, our identity is linked to our IP address. We assume that the ISP is a not-trusted part in this whole system, which means that if our IP is revealed then a malicious part can obtain our real identity by going through the ISP. Hence, the "IP is unknown". If the IP is available to everyone the biggest and most personal part of you online communication has been revealed. Anonymity on the internet in general has to do with keeping our IP address hidden to some extent.

2. The g^x does not need to be encrypted because, to be able to create the same secret key/shared key as Alice will have with Node 1, they would need to first part of the puzzle, which is the g^a, which Alice encrypts with the public key of the Node. If the node were to actually encrypt it, it would need Alice's public key, so only she can decrypt it, but since we don't authenticate Alice before we start the communication, the node never obtains this information.

3. Perfect forward secrecy, means that the information you send today should not be in risk of being exposed if a system is breached in the future. Often this entails using session keys/ephemeral keys, which are short-lived and always renewed. If an attacker is sniffing the connection and manages to obtain one of the keys, the only info they will be able to decrypt, is for the short period this key is valid. All info from the past and all future info will still be secured. It's important because we want to protect the users from potential hacks or key reveals.

4. In a 2-mix chain, can mix 1 learn the identity/IP address of both the sender and the receiver of a message it receives? In a 2-mix chain, the first mix will not learn the identity of both the sender and receiver. It will only learn the identity of the sender, but not the identity of the receiver as that will be available only to the second mix.

5. Using two mixes is almost the same as when we when we prepare the return address to a single mix, but instead we use the multiple (in this case 2) public keys, one for each mixer. The addressee prepares the return message by the public key of both mixes we. They have received an encrypted return address already with the message from the sender. The return address is sent to the addressee as:

   $$K_1(R_1, K_2(R_2, A_x))$$

   where A_x is the encrypted address.

   The receiver will get the encrypted address and send it back through the 2-mix chain by using the public keys of the mixes. The output to the first mix will be in the path will be as follows:

$$K_2(R_2, K_1(R_1, A_x)), R_1(K_x(R_0, M))$$

While the output of the last mix will be as follows:

$$A_x, R_2(R_1(K_x(R_0, M)))$$

6. When Alice creates a Tor circuit, it is the client that chooses the relay nodes, from a downloaded list of relays, these relays then builds the circuit. It depends on relay bandwidth, faster relays have a higher probability of being chosen over slower ones. Tor also tries to not choose relays that are in the same sub-net or belong to the same operator.

7. The purpose of the $H(K_1)$ message sent from the router to Alice is to allow Alice to have a consistency check when she receives the key from the router. She computes the shared key with the router, then hashes it and compares it to the hash she received.

8. She can expect the TCP handshake with the web-server to be performed with low latency because Tor is actively seeking to limit processing delay and bandwidth overhead, which allows them to facilitate anonymous use of interactive real-time applications like browsing, instant messaging or SSH connections.

   Tor also has a standard to use 3 hops between communicating parts. Having less nodes between each client/server will also generate a fast communication. The Tor network also uses symmetric keys between nodes in a circuit instead of asymmetric key, because asymmetric keys are far slower.