# Home Assignment 2A

EITF41, Advanced Web Security

November, 2019

**Solutions to eight A-assignments.**

## A-1

**Motivate why the lecture notes define anonymity roughly as "the IP is unknown".**
The IP-address is the users identity, it's the address of a user on the internet from which data is sent and received. If the IP-address is hidden from all parties on the internet, anonymity is achieved. This mean the IP-address of both communicating parties have to be hidden. Without the IP-address it is impossible to link a certain individual to different actions on the internet. Though usually, only the IP-address is not enough to identify an individual, since the IP-address is received from a ISP (Internet Service Provider). ISP in turn can link an individual to an IP-address. Therefore true anonymity is obtained when the IP-address is undisclosed for ISP, communicating party and other entities on the internet.

## A-2

**Why is it important to have a large volume of traffic in anonymous communication?**
The degree of anonymity is higher if there is a higher anonymity set. That is a large volume of traffic from a large set of anonymous users. This will make analysis more difficult.

## A-3

**Give some examples of data which can be used in traffic analysis.**
Information that can be used is timing, size of packets, amount of packets, etc. With this information, even if the data is encrypted, traffic analysis can sometimes be used to help identify the plain text.

This can be possible when firstly, the transmitted packets are padded to a small set byte-size boundary (if a block cipher is in use). Therefore an eaves dropper can easily learn the approximate length of the original data. Secondly, in interactive mode, every individual keystroke that a user types is sent to the remote machine in a separate IP packet immediately after the key is pressed (except for some meta keys such as Shift or Ctrl).

This property can enable the eavesdropper to learn the exact length of users' passwords. More importantly, the time it takes the operating system to send out the packet after the key press is in general negligible comparing to the inter-keystroke timing. Hence an eavesdropper can learn the precise inter-keystroke timings of users' typing from the arrival times of packets.

## A-5

**What is the purpose of the random value $R_0$ in a Mix?**
$R_0$ is needed in the message in order to prevent an attacker from guessing messages. It is assumed that the attacker can observe all incoming and outgoing messages. If the random string is not used (i.e. only ($K_B$(message)) is sent to B) and an attacker has a good guess that the message message$'$ was sent, he can test whether $K_B$(message$'$) = $K_B$(message) holds, whereby he can learn the content of the message. By appending the random string $R_0$ the attacker is prevented from performing this kind of attack; even if he should guess the correct message (i.e. message$'$ = message is true) he won't learn if he is right since he doesn't know the secret value $R_0$. Practically, $R_0$ functions as a salt.

## A-10

**When using 2 mixes and an untraceable return address, show how the addressee prepares the return message to the original sender.**
When replying to a message, but still keeping the address of the recipient a secret, an encrypted return address is sent in addition to the encrypted message. The return address, $A_x$, is encrypted by the public key of the Mix, after adding some randomness $R_1$. Also in order to allow the reply to be encrypted, a temporary public key, $K_x$, is sent to the recipient. Thus the message $K_1(R_1, A_x)$, $K_x$ is sent to the recipient.

Upon reply, A adds randomness $R_0$ to the return message and encrypts it with the public key $K_x$, $K_x(R_0, \text{M})$. The purpose of $R_0$ is here the same as when sending the original message. This is returned to the Mix together with the encrypted return address. The Mix decrypts the return address with its private key, reads the address $A_x$ and forwards the return message to the original sender S. However, before forwarding the return message it is encrypted using $R_1$ as key. Thus, the randomness $R_1$ serves two purposes in this scheme

- It prevents adversaries, in particular the addressee A, from simply guessing the return address and verify the guess using the public key of the Mix.

- By encrypting the return message with $R_1$, it prevents an eavesdropper to compare the messages entering and leaving the Mix. Without this encryption, $K_x(R_0, \mathrm{M})$, would be seen on both sides of the Mix.

The return address is sent to the addressee as:
$K_1(R_1, K_2(R_2, A_x))\cdots))$.
This encrypted return address is sent in the reply together with the message $K_x(R_0, \mathrm{M})$.
The first Mix in the return path, i.e., the Mix with public key $K_1$, decrypts the outmost encryption layer, removes the randomness $R_1$ and encrypts the reply message with $R_1$.
Thus the output of the first Mix is:
$K_2(R_2, \ldots, K_1(R_1, K_2(R_2, A_x))\cdots)$, $R_1(K_x(R_0, \mathrm{M}))$.
The output of the last Mix, i.e., the message returned to the original sender is given by:
$A_x$, $R_2(R_1(R_0(\ldots R_2(R_1(K_x(R_0, \mathrm{M})))\cdots))$.

## A-13

**It is straightforward to generalize the N - 1 attack to an N - k attack, 0 ¡ k ¡ N. Describe the N - k attack.**
N - k attacks are the most powerful attacks against Mix today. The idea is to reduce the anonymity set (the anonymity degree) for a user. If the adversary can control the generation of several messages that are input to the mix, then she also knows the recipients of these messages. Taking this to the extreme, assume that the adversary controls all senders but one. Then it is easy for the adversary to link the sender and receiver of the remaining message which is done through a traffic analysis.

## A-18

**Why is onion routing called onion routing?**
Onion routing is a technique for anonymous communication over a computer network. In an onion network, messages are encapsulated in layers of encryption, analogous to layers of an onion. The encrypted data is transmitted through a series of network nodes called onion routers, each of which "peels" away a single layer, uncovering the data's next destination. When the final layer is decrypted, the message arrives at its destination.

The encryption is added in layers, just like layers of an onion. In the decryption process, layers are removed one by one.

## A-22

**If you are using the Tor network for you own communication, would you be more or less safe if you would participate as a relay for others in the network as well?**
Users running their own Tor node are more anonymous than users that do not. This is because they generate traffic that is not only their own traffic. Since all packets are

encrypted, an adversary has no way of knowing if cells leaving the node are just relayed from another cell or actually origins in the cell.