COM-301 Computer Security
Exercise 10: Privacy

Privacy

1. Are the following statements True or False:

   (a) Privacy technologies based on using access control to control who access what, protect you from the service provider

   (b) Anonymous communications based on onion routing only protect from adversaries that cannot see both ends of the anonymous communication networks

   (c) Encryption is the only protection needed to ensure privacy online

   (d) Anonymous communications are the only protection needed to ensure privacy online

   (e) An internet without routing security (e.g., enabling BGP hijacking) can break the protection given by low-latency anonymous communication systems

   (f) Using a VPN provides privacy from a local adversary that can only observe the connection from the user to the proxy but not from a global adversary that can observe any link in the network

   Solution:

   (a) False, access control is implemented by the service provider, who enforces some access control policy. To implement this access control the provider does have access to the data, ie., there is no privacy towards the provider.

   (b) True, onion routing is a technique for low-latency anonymous communications networks. As such, it preserves patterns which can be recognized if an adversary can see traffic coming into the network and leaving the network.

   (c) False, encryption only protects content. Privacy online requires protecting also meta-data (e.g., identities, location, etc).

(d) False, anonymous communications only provides protection at the network layer. To obtain full privacy one also needs to take care of the application layer (e.g., use encryption).

(e) True, Anonymous communication systems are overlay networks. This means that onion routers are not network routers. They run at the application layer. Therefore by using BGP hijacking an adversary can reroute traffic in such a way that they can see both traffic coming in and leaving the anonymous communication network . As such they can put themselves in the position of launching an attack. (think of the example seen in the class were all traffic from US was routed through Belarus. In this case a Belarus ISP would be a Global adversary for a US citizen visiting a US website would have an adversary)

(f) True, a local adversary can only see the IP of the VPN service. Thus it cannot learn the destination of the traffic. However, a global adversary that can observe any link can see the input and the output to the VPN. A simple timing attack would suffice to track where incoming flows go.

2. We have seen in the class that a problem with the Tor network, and in general any low-latency anonymous communication systems, is that the shape of the traffic enables an adversary to trace flows through the network. A possible option to fight this type of attacks is to "pad" the connection, i.e., insert enough dummy packets so that all connections look similar. Does this approach make sense to:

   (a) Conceal whether a user is visiting a news website (e.g., CNN, BBC, etc) or a small static web such as http://www.guimp.com/

   (b) Conceal which EPFL employee website a user is visiting (e.g., `https://people.epfl.ch/`)

   (c) Conceal which movie you are downloading from the Pirate Bay.

   (d) Conceal when electronic votes are being sent.

   [HINT: think about the bandwidth overhead you would need to protect the traffic]

   Solution:

   (a) Not really. Nowadays news websites contain a fair amount of images, videos, etc. As such, they require to download a fair amount of bytes. In order to conceal the small static webs, one would need to add enough "padding" for the small websites to look like the news websites. This creates a lot of overhead in the network.

   (b) Yes, it would make sense. In the case of the EPFL employees websites, all of them are very similar. Thus, very little padding would make them indistinguishable.

(c) Not really. Similarly to the news websites, movies can be quite large and there may be a large difference between some of the archives in the database (from Megas in a small videoclip, to Gb in an HQ movie). The overhead to make them all similar would be quite large.

(d) Yes, it would make sense. Votes are very small. With little padding, one could fake the sending of "dummy" votes effectively concealing when real votes are being sent.

3. Let us assume that a COM-301 TA decides to invest the money from selling the exam questions into helping the Tor network instead of going to Hawaii. Thus, he buys 1000 machines and sets them up in the basement as onion routers.

   (a) Does this increase the Tor network capabilities to offer anonymity? Against which adversary?

   (b) If instead only one assistant, each of the COM-301 students puts 10 machines in their own basement, would the situation improve?

   [HINT: think about who could see the traffic going through the new machines should they be incorporated to the Tor network.]

   Solution:

   (a) No, it does not. Having many servers increases the number of possible routes, but if all of them are in the same place there is no diversity. An adversary that can see one machine can see all of them (i.e., can see the full Tor circuit and use traffic analysis to find where flows go). The problem of this approach is the little diversity in the nodes that puts certain adversaries in an advantageous position. In this case, for instance the Swiss ISP serving the TA's basement internet. Note that these additional servers do not increase anonymity with respect to an adversary that cannot see them either. For this adversary, it actually makes no difference that there are more or less circuits that can go through the TA's basement. On the other hand, having more nodes does increase the capacity of the network.

   (b) This improves the situation, for instance against the ISP if the students are served by different ISPs. Yet, for instance the Swiss government could have the capability to observe the traffic in all of them (e.g., with a subpoena).

4. We have seen in the class that anonymous credentials enable users to prove possession of an attribute (e.g., Bob is subscribed to the newspaper, Thomas is older than 18, etc.), and also that they are unlinkable across uses (e.g., Bob's visits to the newspaper cannot be linked to each other).

Compare the privacy that Bob has when he visits CNN over Tor, and when he visits CNN without Tor.

Solution:
When Bob does not use Tor, even though anonymous credentials ensure that his visits to the website are not linkable *at the application layer*, his network metadata (e.g., IP) make his actions linkable.

When Bob uses Tor: a) if he creates a new circuit every time, the network metadata is not the same and thus the server cannot link his actions; b) if he reuses the circuit, the server can link connections as coming from the same exit node. However, from the servers' point of view there is no way to know if those come from the same circuit, or from a different circuit with the same exit. Therefore, Bob's actions are still unlinkable.

We did not explain this in the class, but even if Bob's IP would be dynamic, his browser is unique and the information it sends to the web server can be used to link Bob's visits (https://amiunique.org/)

When using Tor, one uses the Tor Browser, which is configured to securely use Tor. This Browser is common to all Tor users and therefore the service provider cannot assume that all visits with the Tor browser are Bob's. Bob's actions are still unlinkable.