Security Protocols and Application — Final Exam
Solution

Philippe Oechslin and Serge Vaudenay

27.6.2016

- duration: 3h00
- no document allowed
- a pocket calculator is allowed
- communication devices are not allowed
- the exam invigilators will not answer any technical question during the exam
- the answers to each exercise must be provided on separate sheets
- readability and style of writing will be part of the grade
- do not forget to put your name on every sheet!

The exam grade follows a linear scale. In each exercise, each question has the same weight. Both exercises have the same weight.

1 Signal Protocol, the Improvement over OTR

Q.1 The words "axolotl" and "ratchet" both have double meanings. They have a meaning outside of computer science and they have a meaning in cryptography.
Answer the following questions.
- Give the two meanings of "axolotl". How is it related to this exercise?
- Give the two meanings of "ratchet". How is it related to this exercise?
- What does "OTR" mean?
- Name at least six objectives of Signal.

An axolotl is a threatened species of Mexican salamander whose name comes from a Mexican god with a dog head. It is also the former name of the Signal protocol, the topic of this exercise.
A ratchet is a mechanical device which can only move forward. For instance, this is a wheel which can only turn in one direction due to a latch preventing to turn backward. In cryptography, it is a scheme to update key materials in a one-way direction to ensure forward secrecy. Keys are updated and old keys are destroyed to avoid having long term keys which could leak. Indeed, we want to avoid that disclosing a long term key compromises privacy (forward secrecy or future secrecy).
OTR refers to the Off-The-Record messaging protocol. It is a protocol which was design before Signal for instant messaging protocols.
For Signal, some objectives are:
- *secure messaging (confidentiality, authenticity, integrity of messages),*
- *forward and future secrecy (confidentiality preserved even though secret leaks),*
- *deniability (no transferable proof of message authorship leaks),*
- *being asynchronous (can be done offline),*
- *detect replay/reorder/deletion attacks,*
- *allow decryption of out-of-order messages,*
- *and don't leak metadata.*

Q.2 In the Signal Protocol, when Alice wants to send a message to Bob for the first time, she first fetches Bob's long-term public key bG and one of his one-time public prekeys $x_{b,i}G$ from a central server (then, the server erases this prekey). All public keys are in Curve25519. Then, she selects her ephemeral public key $x_{a,\mathsf{eph}}G$ and she sets $x_{b,\mathsf{eph}}G = x_{b,i}G$. In other cases, when Alice wants to send a message to Bob, she already has a state containing $bG, x_{b,\mathsf{eph}}G, x_{a,\mathsf{eph}}G$. To send a message to Bob, Alice takes her long-term public key aG then computes offline

$$\mathsf{secret} = ax_{b,\mathsf{eph}}G \| bx_{a,\mathsf{eph}}G \| x_{a,\mathsf{eph}}x_{b,\mathsf{eph}}G$$

Alice's message to Bob is then secured with secret. Together with her message, she advertises (sends) her ephemeral public key $x_{a,\mathsf{eph}}G$ and identifies which of Bob's prekey was used as well. So, Bob will be able to build a state containing $aG, x_{a,\mathsf{eph}}G, x_{b,\mathsf{eph}}G$.
Answer the following questions.

- What is Curve25519?
- What are the secret keys and where do they belong to?
- When (and why) does Alice erase her ephemeral key?
- Why don't we use abG instead of these three values?

- Curve25519 *is an elliptic curve. We use a subgroup generated by some base point G of order $q = 2^{252} + 27742317777372353535851937790883648493$. There is a mapping such that each 256-bit string corresponds to a point in the elliptic curve.*
- *Secret keys are integers modulo q. For instance, the secret key of x_aG is x_a.*
- *Alice erases her ephemeral key when she receives a reply from Bob. Her ephemeral key will be used by Bob to send a reply and step the ratchet. So, she needs it to compute the next secret.*
- *If we used abG, an adversary could replay (assuming that Bob does not realize he already has a state for Alice). Another problem is that if in the future Bob's key b leaks, then this message abG could be used to see that Alice initialized a communication with Bob and could also be used to derive* secret. *By using ephemeral keys which are erased, we have forward/future secrecy.*

Q.3 The Signal Protocol uses the Double Ratchet Algorithm. The ratchet key is RK. When Alice sends her first message to Bob, the 3DH handshake is used to derive RK from secret. Then, for when Alice or Bob replies with a new message, RK is updated by

$$\mathsf{RK} \| \mathsf{NHK} \| \mathsf{CK} \leftarrow \mathsf{KDF}(\mathsf{HMAC}(\mathsf{RK}, x_{a,\mathsf{eph}}x_{b,\mathsf{eph}}G))$$

This way, a chain keys CK is derived. Then, a sequence of message keys MK are derived by iterative hashing on the chain key. A message is encrypted with MK.
In addition to making CK, RK is also used to derive a header key HK which is used to encrypt the ephemeral key $x_{\cdot,\mathsf{eph}}G$, the message number Ns within the ratchet, and the number PNs of sent messages in the previous ratchet. The NHK is the next value of HK in the next ratchet. So, Alice sends

$$\mathsf{Enc}_{\mathsf{HK}}(\mathsf{Ns}, \mathsf{PNs}, x_{a,\mathsf{eph}}G) \| \mathsf{Enc}_{\mathsf{MK}}(\mathsf{plaintext})$$

Encryption follows the encrypt-then-MAC principle using AES256 and HMAC-SHA256.
Answer the following questions.

- Explain why we call this algorithm “Double Ratchet” and why it is done like this.
- Why do we send Ns and PNs?
- Why don’t we use NHK right away instead of waiting for the next ratchet?
- Could the MAC of messages be used as a signature from Alice to prove that she did send a message to Bob?

- *There is one ratchet for the sequence of* RK *and another ratchet for the sequence of* MK*. This allows to continue to send messages with new keys even though no response was received. Indeed,* RK *is renewed at every round trip communication. But Alice may send several messages to Bob before Bob answers.*
  The ratchet for RK *is inherited from OTR. The ratchet for* MK *is inherited from SCIMP, the Silent Circle Instant Messaging Protocol (IM protocol from the Silent Circle company).*
  It is hard to use the SCIMP-style ratchet as the main one without assuming Alice and Bob to be well synchronized. Indeed, if Alice and Bob sends a message at the same time, it is not clear how to step the ratchet in this case. If we impose round trips to step the ratchet, it means that either Alice cannot send several messages or cannot step the ratchet if she does it.
- *The values* Ns *and* PNs *are used by the receiver to reconstruct the order or messages when they arrive. By having* PNs*, the receiver can check that all messages from the last ratchet have been received. If all are received, he can erase the keys which are no longer useful. By having* Ns*, he can further deduce which* MK *to derive to decrypt the message.*
- *We cannot use* NHK *right away because it is used to encrypt* $x_{\cdot,\mathsf{eph}}$ *and the receiver needs it to derive* HK*.*
- *The MAC is computed with a key which is derived from ephemeral keys. So, no link to Alice of Bob can be proven. The message is deniable.*

2 Powerspy

Q.1 The powerspy attack is used to try to infer a phone's location without using the sensors supposed to provide location information.
Answer the following questions.
- Which sensors are usually used to provide location information to applications ?
- Why would an application not want to use these sensors ?

> - *The GPS sensor and the Wifi BSSID can both be used to locate a phone. When the signal of the GPS satellites are strong enough the GPS sensor can give a accurate location. The location of a Wifi network can be looked up in public and private databases based on the unique BSSID.*
> - *The user may not give permission to the application to use localisation sensors. If the application wants to know the localisation without having this permission, it needs to extract location information from another sensor.*

Q.2 The powerspy attack is able to deduce the location of the phone from its power consumption.
Answer the following questions.
- What makes the power consumption dependent on the location of the phone ?
- Given a single measurement of the (electric) current consumption of a phone, what can be said about its location ? Explain in a few words.
- Cite three goals concerning location that can be achieved by constantly monitoring the phone's current consumption.

> - *The power consumption of the radio hardware depends, among others, of the distance to the cell tower to which the phone is connected.*
> - *Nothing. The current consumption depends on many other parameters of the phone. A single measurement is not enough to reveal information about location.*
> - *Route distinguishability, real-time tracking and inference of new routes.*

Q.3 Suppose you want to use Powerspy to identify which route among a set of possible routes you have taken to come to the exam this morning.
Answer the following questions.
- What work did you have to do in advance of the exam to be able to carry out the attack today ?
- What preprocessing do yo have to apply to the measurements in order to facilitate the identification of the route ?
- Describe in a few sentences the algorithm that leads from the preprocessed measurements of this morning's power consumption to the identification of the route you took.

> - *We need to record the power consumption profiles along each route.*
> - *We normalize the measurements by subtracting their mean value and dividing by the standard deviation.*
> - *We use a matching algorithm like DTW which can compress or stretch the time series when matching them in order to take into account differences in speed. The best match is then considered as the identified route.*

Q.4 Explain why the power consumption profile is different depending on the direction in which a route is taken.

The power consumption changes when a phone disconnects from one base station to connect to the next one. This handover is only done when the signal of the new base station is better that the one of the current one. The difference of signal necessary for the phone to switch (the hysteresis) places the point of the handover point further down the road in the direction the phone is moving. This point thus depends on the direction in which the phone is moving.

Q.5 Different possibilities exist to prevent this attack.

Answer the following questions.

- What mechanism does an application use to know the current power consumption ?
- Why is adding noise to the available measurement of the power consumption not a satisfying solution ?
- What would be a simple and logical solution to prevent the attack ?

- *The current consumption can simply be read from a file (/sys/class/powersupply/battery/current_now).*
- *The matching algorithm already eliminates noise by downsampling the measures. The added noise would have to be very important or very specific.*
- *The simplest would be to handle this file like other sensors and require a specific permission to access it.*

Q.6 Finally, can you find a useful and honest usage of the power consumption for the localisation of a phone ?

For example:
- *As an aid for car navigation when the GPS signal is bad due to buildings or tunnels*
- *As a way to do localisation with a lower power consumption. Eg. as long as the power consumption indicates that we are on the right route, no need to activate GPS or Wifi.*