Description of the Needham Schroeder public key protocol and its attack

The Needham Schroeder public key protocol can be described as follows.

$$\begin{array}{lll} A & \to B: & \{A, N_a\}_{\mathsf{pub}(B)} \\ B & \to A: & \{N_a, N_b\}_{\mathsf{pub}(A)} \\ A & \to B: & \{N_b\}_{\mathsf{pub}(B)} \end{array}$$

Initial knowledge: We suppose that agents A and B initially know public keys $\mathsf{pub}(C)$ of agent C, for any agent C.

Data generated during the protocol: N_a is a nonce generated by A. N_b is a nonce generated by B.

Protocol description: Alice starts the protocol by sending her identity A together with a freshly generated random number N_a. This message is encrypted using an asymmetric encryption algorithm with B's public key (denoted $\mathsf{pub}(B)$). We suppose that only agent Bob (whose identity is B) knows the secret key corresponding to $\mathsf{pub}(B)$.
Next Bob receives the message $\{A, N_a\}_{\mathsf{pub}(B)}$ sent by Alice. Using his private key, Bob decrypts the message. He sends the received nonce N_a together with a freshly generated nonce N_b encrypted with A's public key ($\mathsf{pub}(A)$) to Alice.
Finally Alice receives the message $\{N_a, N_b\}_{\mathsf{pub}(A)}$. She decrypts the message and checks that the nonce N_A corresponds to the nonce previously generated and sent to Bob. She sends the nonce N_b to Bob encrypted with Bob's public key. Upon reception of this message Bob decrypts it and checks that the nonce corresponds to the one previously generated.

Security properties:

- *Authentification:* When Bob receives the last message ($\{N_b\}_{\mathsf{pub}(B)}$), this message was indeed sent by Alice.
- *Confidentiality:* Both Alice and Bob are the only ones to know N_b.

Cost: $53 + 53 + 3 = 109$

- first message: $1 + (50 + 1 + 1) + 1 = 53$
- second message: $1 + (50 + 1 + 1) + 1 = 53$
- third message: $1 + 1 + 1 = 3$

Attack on the Needham-Schroeder protocol

17 years after the publication of the protocol Gavin Lowe discovered an attack that may occur in the presence of an active adversary. The attack has been coined *man-in-the-middle attack*. It is illustrated in figure 1. Agent A starts a session with a dishonest agent C. Agent C uses this message to fake being A to B. B responds to A. As B's message contains the nonce N_a, A accepts the message thinking it originates from C. Therefore A sends to C the nonce N_b encrypted with C's public key. C can recover the nonce N_b and end the protocol with B who thinks having executed the protocol with B.

$$\text{A} \xrightarrow{\{A, N_a\}_{\mathsf{pub}(C)}} \text{C}$$
$$\downarrow$$
$$\text{C(A)} \xrightarrow{\{A, N_a\}_{\mathsf{pub}(B)}} \text{B}$$
$$\text{C(A)} \xleftarrow{\{N_a, N_b\}_{\mathsf{pub}(A)}} \text{B}$$
$$\downarrow$$
$$\text{A} \xleftarrow{\{N_a, N_b\}_{\mathsf{pub}(A)}} \text{C}$$
$$\text{A} \xrightarrow{\{N_b\}_{\mathsf{pub}(C)}} \text{C}$$
$$\downarrow$$
$$\text{C(A)} \xrightarrow{\{N_b\}_{\mathsf{pub}(B)}} \text{B}$$

Figure 1: Lowe's attack on the Needham Schroeder public key protocol.

Game designed by V. Cortier. Translation kindly provided by S. Kremer.