OTR Properties

Joe Bonneau
Andrew Morrison

February 13, 2006

1 Introduction

In order to perform a formal security analysis of the Off-The-Record messaging protocol we shall first define the properties we wish to verify. With such properties defined, we shall model OTR using the Murϕ model checker.This is because we need additional flexibility to model perfrect forward secrecy and weak deniability that model checkes like `AVISPA` and `PRISM` do not offer. Throughout, we will refer to the principles Alice (A), Bob (B), Mallory (M), Alice$'$ (A'), Bob$'$ (B'), and Justin (J).

For the conventional properties of secrecy, perfect forward secrecy, authentication, and message integrity, Alice and Bob are honest agents wishing to hold an OTR conversation such that all of the soon to be defined properties hold. Mallory is a malicious agent with full control of the network looking to break any property she is capable of.

For the deniability properties, the informant Alice$'$ represents an agent whom Bob trusts and will carry out an honest conversation with. However, Alice$'$ is attempting to prove the contents of the conversation to Justin. Justin can provide information to Alice$'$ before or during a conversation, and can receive all of the information available to Alice$'$, but Justin cannot trust Alice$'$ or any information provided by Alice$'$. This represents the real-world scenario of two complicit criminals Alice and Bob, one of who decides to turn the other into the justice system. The justice system, represented by Justin, cannot trust that Alice$'$ is honest, it could be that Alice$'$ is attempting to frame Bob. Bob$'$ is defined symmetrically to Alice$'$.

2 Security Properties

The OTR protocol aims to provide a system of digital communication resembling a casual private conversation in which the honest principles involved may be assured *secrecy*, *authentication* and more interestingly *perfect forward secrecy* and *deniability*.

2.1 Secrecy

A secrecy invariant shall be defined such that for a conversation between principles $A^* \in \{A, A'\}$ and $B^* \in \{B, B'\}$, no other agent shall possess the tuple $\{\text{AES}_{e_{k_t}}(msg), e_{k_t}, c_t\}$ for any time t.

2.2 Perfect Forward Secrecy

Perfect forward secrecy shall be defined such that for a conversation between principles $A^* \in \{A, A'\}$ and $B^* \in \{B, B'\}$, no other agent in possession of the tuple $\{\text{AES}_{e_{k_t},c_t}(msg), e_{k_t}, c_t\}$ at time t shall be able to to learn any information about the same tuple for time $t' < t$. To model this we will add an action for Mallory which gives her all of Alice or Bob's current secret keys at any point in the conversation. This is the equivalent of Mallory hacking into one of the honest principal's computers midway through a conversation. This should not enable her to decrypt past messages which she has intercepted and stored.

2.3 Authentication

An initial conversation in which an untrusted public key is transmitted between principles is trivially insecure and shall be treated as a special case. For subsequent conversations, we shall treat a stored pair $\{p, \text{pub}_p\}$ for principle p as trusted.

Authentication on post initial key exchange shall be defined such that for any initiator wishing to communicate with principle p in receipt of $\text{sig}_p(M_p)$ from responder r, it must be the case that $r = p$.

2.4 Integrity

For $X^t_{A^*} = (\text{keyid}^t_A, \text{keyid}^t_{B^*}, \text{next\_dh}_t, c_t, \text{AES}_{e_{k_t},c_t}(msg))$ transmitted at time t and integrity provided by $\text{MAC}_{m_{k_t}}(X^t_{A^*})$, no agent $M \notin \{A^*, B^*\}$ shall be capable of producing a valid pair $\{Y^t_{A^*}, \text{MAC}_{m_{k_t}}(Y^t_{A^*})\}$ for $Y^t_{A^*} \neq X^t_{A^*}$. It should be noted that the integrity of $X^t_{A^*}$ shall be made explicitly untrusted at time $t + 1$ due to the intentional publication of MAC keys m_{k_t}.

2.5 Plausible Deniability

We define *plausible deniability* via two categories; *weak deniability* in which it may be proven that both A and B have all necessary key material to produce any given message and *strong deniability* in which it may be proven that principles other than A and B are capable of producing valid messages.

2.5.1 Weak Deniability

During data exchange, the ability to transmit a valid message requires $keyid_A$, $keyid_B$, $next\_dh, t, K_e$ and t. Weak deniability shall be defined as the property that both A^* and B^* possess the full set of necessary numbers. Given two parties with

the necessary numbers, it cannot be proven that either one was the legitimate author.

2.5.2 Strong Deniability

We define *strong deniability* as the ability to claim that not only could A and B have created a given message, but anyone could have created a valid message. This property may be modeled by the ability of an outside agent M who upon being given a valid transcript of a conversation may produce a different yet still cryptographically valid transcript.