KEY DISTRIBUTION: PKI and SESSION-KEY EXCHANGE

The public key setting

$$
\begin{array}{ccc}
\text{Alice} & & \text{Bob}^{pk[A]} \\[1em]
M \leftarrow \mathcal{D}_{sk[A]}(C) & \xleftarrow{\quad C \quad} & C \stackrel{\$}{\leftarrow} \mathcal{E}_{pk[A]}(M) \\[1em]
\sigma \stackrel{\$}{\leftarrow} \mathcal{S}_{sk[A]}(M) & \xrightarrow{\quad M,\sigma \quad} & \mathcal{V}_{pk[A]}(M,\sigma)
\end{array}
$$

Bob can:

- send encrypted data to Alice
- verify her signatures

as long as he has Alice's public key $pk[A]$.

But how does he get $pk[A]$?

Distributing public keys

How about:

$$
\begin{array}{ccc}
\text{Alice} & & \text{Bob} \\
(pk[A], sk[A]) \xleftarrow{\$} \mathcal{K} & \xrightarrow{\quad Alice,\, pk[A] \quad} & \\
M \leftarrow \mathcal{D}_{sk[A]}(C) & \xleftarrow{\quad C \quad} & C \xleftarrow{\$} \mathcal{E}_{pk[A]}(M) \\
\sigma \xleftarrow{\$} \mathcal{S}_{sk[A]}(M) & \xrightarrow{\quad M, \sigma \quad} & \mathcal{V}_{pk[A]}(M, \sigma)
\end{array}
$$

Man-in-the-middle attack

$$\begin{array}{ccc}
\text{Adversary } E & & \text{Bob} \\
(pk[E], sk[E]) \xleftarrow{\$} \mathcal{K} & \xrightarrow{\quad Alice,\, pk[E] \quad} & \\
M \leftarrow \mathcal{D}_{sk[E]}(C) & \xleftarrow{\quad C \quad} & C \xleftarrow{\$} \mathcal{E}_{pk[E]}(M) \\
\sigma \xleftarrow{\$} \mathcal{S}_{sk[E]}(M) & \xrightarrow{\quad M, \sigma \quad} & \mathcal{V}_{pk[E]}(M, \sigma)
\end{array}$$

Messages that Bob encrypts to Alice are obtained by E, and E can forge Alice's signatures.

The authenticity problem and PKI

Bob needs an authentic copy of Alice's public key.

The PKI (Public Key Infrastructure) is responsible for ensuring this.

Usually it is done via certificates.

Certificate Process

- Alice generates *pk* and sends it to CA
- CA does identity check
- Alice proves knowledge of secret key to CA
- CA issues certificate to Alice
- Alice sends certificate to Bob
- Bob verifies certificate and extracts Alice's *pk*

Generate key and send to CA

Key generation: Alice generates her keys locally via $(pk, sk) \xleftarrow{\$} \mathcal{K}$

Send to CA: Alice sends $(Alice, pk)$ to a certificate authority (CA).

Identity check

Upon receiving $(Alice, pk)$ the CA performs some checks to ensure pk is really Alice's key:

- Call Alice by phone
- Check documents

These checks are out-of-band.

Proof of knowledge

The CA might have Alice sign or decrypt something under *pk* to ensure that Alice knows the corresponding secret key *sk*.

This ensures Alice has not copied someone else's key.

Certificate Issuance

Once CA is convinced that *pk* belongs to Alice it forms a certificate

$$CERT_A = (CERTDATA, \sigma),$$

where σ is the CA's signature on *CERTDATA*, computed under the CA's secret key $sk[CA]$.

CERTDATA:

- *pk*, *ID* (Alice)
- Name of CA
- Expiry date of certificate
- Restrictions
- Security level
- ...

The certificate $CERT_A$ is returned to Alice.

Certificate usage

Alice can send $CERT_A$ to Bob who will:

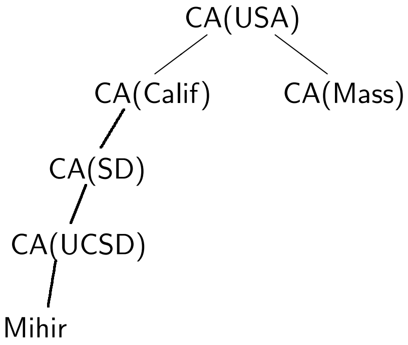
- $(CERTDATA, \sigma) \leftarrow CERT_A$
- Check $\mathcal{V}_{pk[CA]}(CERTDATA, \sigma) = 1$ where $pk[CA]$ is CA's public key
- $(pk, Alice, expiry, \ldots) \leftarrow CERTDATA$
- Check certificate has not expired
- $\ldots$

If all is well we are ready for usage.

How does Bob get $pk[CA]$?

CA public keys are embedded in software such as your browser.

Certificate hierarchies



$\underline{CERT_{Mihir}}$

$CERT[CA(USA) : CA(Calif)]$
$CERT[CA(Calif) : CA(SD)]$
$CERT[CA(SD) : CA(UCSD)]$
$CERT[CA(UCSD) : Mihir]$

$$CERT[X : Y] = (pk[Y], Y, \ldots, \mathcal{S}_{sk[X]}(pk[Y], Y, \ldots))$$

To verify $CERT_{Mihir}$ you need only $pk_{CA[USA]}$.

Why certificate hierarchies?

- It is easier for CA(UCSD) to check Mihir's identity (and issue a certificate) than for CA(USA) since Mihir is on UCSD's payroll and UCSD already has a lot of information about him.
- Spreads the identity-check and certification job to reduce work for individual CAs
- Browsers need to have fewer embedded public keys. (Only root CA public keys needed.)

Revocation

Suppose Alice wishes to revoke her certificate $CERT_A$, perhaps because her secret key was compromised.

- Alice sends $CERT_A$ and revocation request to CA
- CA checks that request comes from Alice
- CA marks $CERT_A$ as revoked

Certificate revocation lists (CRLs)

CA maintains a CRL with entries of form

$$(CERT, \text{Revocation date})$$

This list is disseminated.

Before Bob trusts Alice's certificate he should ensure it is not on the CRL.

Revocation Issues

- November 22: Alice's secret key compromised
- November 24: Alice's $CERT_A$ revoked
- November 25: Bob sees CRL

In the period Nov. 22-25, $CERT_A$ might be used and Bob might be accepting as authentic signatures that are really the adversary's. Also Bob might be encrypting data for Alice which the adversary can decrypt.

Revocation in practice

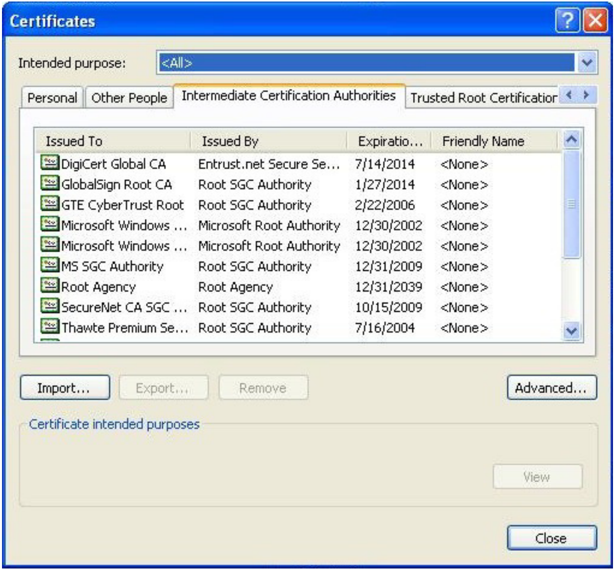
- VeriSign estimates that 20% of certificates are revoked
- In practice, CRLs are huge

Revocation is a big problem and one of the things that is holding up widespread deployment of a PKI and use of public-key cryptography.

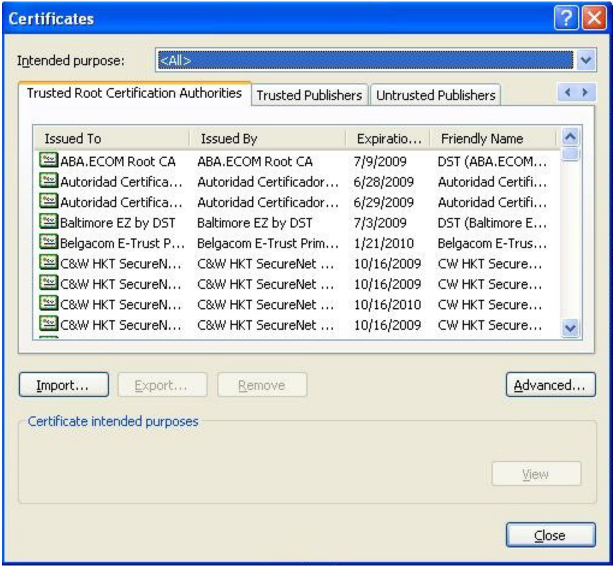
PGP

In PGP, there are no CAs. You get Alice's public key from Carol and decide to what extent you want to trust it based on your feelings about Carol. Requires user involvement.

Certificate Examples

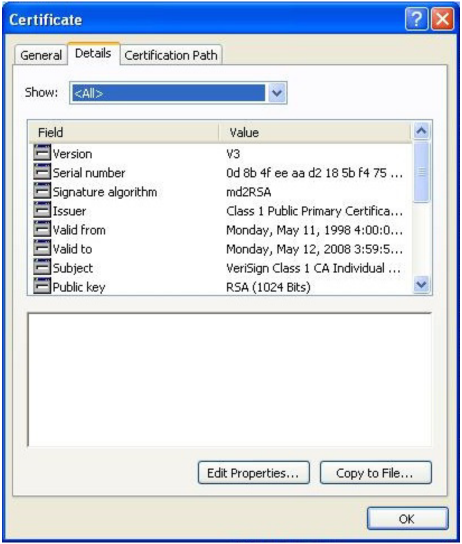


Certificate Examples

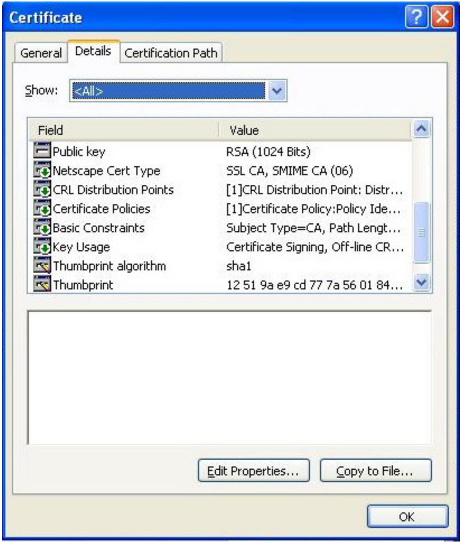


Certificate Examples

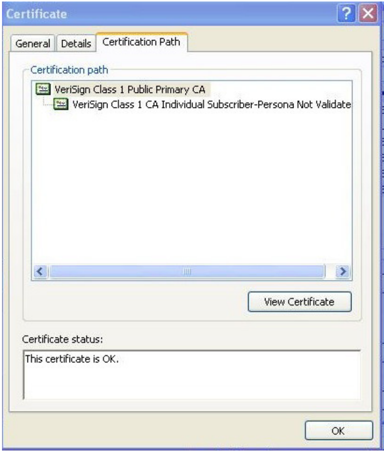
Certificate Examples



Certificate Examples



Certificate Examples



Shared key setting

$$
\begin{array}{ccc}
\text{Alice}^K & & \text{Bob}^K \\[1em]
M \leftarrow \mathcal{D}_K(C) & \xleftarrow{\quad\quad C \quad\quad} & C \xleftarrow{\$} \mathcal{E}_K(M) \\[1em]
\sigma \xleftarrow{\$} \mathcal{T}_K(M) & \xrightarrow{\quad\quad M,\sigma \quad\quad} & \mathcal{V}_K(M,\sigma)
\end{array}
$$

Alice and Bob can

- send each other encrypted data
- verify each other's MACs

Can be preferable to public key setting because computation costs are lower.

But how do Alice and Bob get a shared key?

Diffie-Hellman Key Exchange

Let $G = \langle g \rangle$ be a cyclic group of order m and assume G, g, m are public quantities.

Alice		Bob
$x \xleftarrow{\$} \mathbf{Z}_m;\ X \leftarrow g^x$	$\xrightarrow{\text{Alice}, X}$	
$K \leftarrow Y^x$	$\xleftarrow{\text{Bob}, Y}$	$y \xleftarrow{\$} \mathbf{Z}_m;\ Y \leftarrow g^y$
		$K \leftarrow X^y$

$$Y^x = (g^y)^x = \underbrace{g^{xy}}_{K} = (g^x)^y = X^y$$

This enables Alice and Bob to agree on a common key K which can subsequently be used, say to encrypt:

Alice		Bob
$M \leftarrow \mathcal{D}_K(C)$	$\xleftarrow{C}$	$C \xleftarrow{\$} \mathcal{E}_K(M)$

Security of DH Key Exchange under Passive Attack

$$
\begin{array}{rcl}
\text{Alice} & & \text{Bob} \\
x \stackrel{\$}{\leftarrow} \mathbf{Z}_m;\ X \leftarrow g^x & \xrightarrow{\text{Alice}, X} & \\
K \leftarrow Y^x & \xleftarrow{\text{Bob}, Y} & y \stackrel{\$}{\leftarrow} \mathbf{Z}_m;\ Y \leftarrow g^y \\
 & & K \leftarrow X^y
\end{array}
$$

Eavesdropping adversary gets $X = g^x$ and $Y = g^y$ and wants to compute g^{xy}. But this is the (presumed hard) CDH problem.

Conclusion: DH key exchange is secure against passive (eavesdropping) attack.

Security of DH Key Exchange under Active Attack

Man-in-the-middle attack:

$$
\begin{array}{rcl}
E & & \text{Bob} \\
x \stackrel{\$}{\leftarrow} \mathbf{Z}_m;\ X \leftarrow g^x & \xrightarrow{\text{Alice}, X} & \\
K \leftarrow Y^x & \xleftarrow{\text{Bob}, Y} & y \stackrel{\$}{\leftarrow} \mathbf{Z}_m;\ Y \leftarrow g^y \\
 & & K \leftarrow X^y \\
M \leftarrow \mathcal{D}_K(C) & \xleftarrow{C} & C \stackrel{\$}{\leftarrow} \mathcal{E}_K(M)
\end{array}
$$

Adversary E impersonates Alice so that:

- Bob thinks he shares K with Alice but E has K
- E can now decrypt ciphertexts Bob intends for Alice

Conclusion: DH key exchange is insecure against active attack

When is key agreement possible?

In the presence of an active adversary, it is impossible for Alice and Bob to

- start from scratch, and
- exchange messages to get a common key unknown to the adversary

Why? Because there is no way for Bob to distinguish Alice from the adversary.

Alice and Bob need some a priori “information advantage” over the adversary. This typically takes the form of long-lived keys.

Settings and long-lived keys

- Public key setting: A has pk_B and B has pk_A
- Symmetric setting: A, B share a key K
- Three party setting: A, B each share a key with a trusted server S.

These keys constitute the long-lived information.

Session keys: The "real" key distribution problem

In practice, Alice and Bob will engage in multiple communication "sessions." For each, they

- First use a session-key distribution protocol, based on their long-lived keys, to get a fresh, authentic session key;
- Then encrypt or authenticate data under this session key for the duration of the session

Why session keys?

- In public-key setting, efficient cryptography compared to direct use of long-lived keys
- Security attributes, in particular enabling different applications to use keys in different ways and not compromise security of other applications

Session key distribution

- Hundreds of protocols
- Dozens of security requirements
- Lots of broken protocols
- Protocols easy to specify and hard to get right
- Used ubiquitously: SSL, TLS, SSH, ...

Security of Session Key

Question: What is desired security attribute of session key?

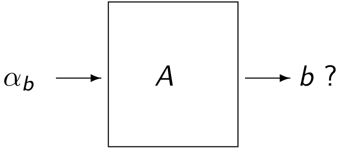
Security of Session Key

Question: What is desired security attribute of session key?

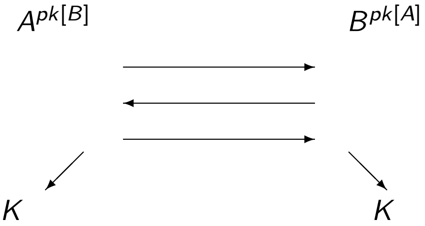
Answer: It should be indistinguishable from random to adversary

At end of protocol:

$$b \xleftarrow{\$} \{0,1\};\ \alpha_0 \leftarrow \alpha;\ \alpha_1 \xleftarrow{\$} \{0,1\}^{|\alpha|}$$



Session key exchange in public key setting



Most important type of session key exchange in practice, used in all communication security protocols: SSL, SSH, TLS, IPSEC, 802.11, ...

Protocol KE1: Basic Exchange

$$
\begin{array}{lcl}
A^{pk[B]} & & B^{pk[A]} \\
 & \xrightarrow{\quad A,\ R_A \quad} & \\
 & \xleftarrow{\quad R_B,\ C,\ B,\ \mathrm{Sign}_B(A, B, R_A, R_B, C) \quad} & C \xleftarrow{\$} \mathcal{E}_A(K) \\
 & \xrightarrow{\quad A,\ \mathrm{Sign}_A(A, B, R_A, R_B) \quad} &
\end{array}
$$

Session key K is chosen by B.

R_A, R_B are random nonces.

$\mathrm{Sign}_P(M)$ is P's signature of M under $sk[P]$. It is verifiable given $pk[P]$.

$\mathcal{E}_A(\cdot)$ is encryption under A's public key $pk[A]$, decryptable using $sk[A]$.

Forward secrecy

$A^{pk[B]}$ $\qquad\qquad\qquad\qquad$ $B^{pk[A]}$

$\xrightarrow{A,\, R_A}$

$\xleftarrow{R_B,\, C,\, B,\, \mathrm{Sign}_B(A, B, R_A, R_B, C)}$ $\qquad C \xleftarrow{\$} \mathcal{E}_A(K)$

$\xrightarrow{A,\, \mathrm{Sign}_A(A, B, R_A, R_B)}$

$\xleftarrow{C_B}$ $\qquad C_B \xleftarrow{\$} \mathcal{E}_K(M)$

Nov. 20: Adversary E records above flows.
Dec. 18: A's, system compromised and $sk[A]$ exposed.
Dec. 19: A revokes $pk[A]$ so that no further damage is done but cannot prevent E from $K \leftarrow \mathcal{D}_{sk[A]}(C)$; $M \leftarrow \mathcal{D}_K(C_B)$.

Can we achieve forward secrecy: Privacy of communication done prior to exposure of $sk[A]$ is not compromised?

KE2: Adding forward secrecy

$A^{pk[B]}$ $B^{pk[A]}$

$$A \xrightarrow{\quad A, g^a \quad} B$$

$$A \xleftarrow{\quad g^b, B, \mathrm{Sign}_B(A, B, g^a, g^b) \quad} B$$

$$A \xrightarrow{\quad A, \mathrm{Sign}_A(A, B, g^a, g^b) \quad} B$$

Session key is $K = H(\boxed{A, B, g^a, g^b}, g^{ab})$.

Adversary E records above flows on Nov. 20. On Dec. 18, $sk[A]$ is exposed. This allows E to forge A's signatures, but A can address this by revoking $pk[A]$. But $sk[A]$ does not help E obtain K.

There is no public-key encryption here, only signatures.

All standard protocols use DH to get forward security in ways like this.

Passwords

A password is a human-memorizable key.

Attackers are capable of forming a set D of possible passwords called a dictionary such that

- If the target password pw is in D and
- The attacker knows $\overline{pw} = f(pw)$, the image of pw under some public function f.

then the target password can be found via

$$\begin{array}{l} \text{for all } pw' \in D \text{ do} \\ \quad \text{if } f(pw') = \overline{pw} \text{ then return } pw' \end{array}$$

This is called a dictionary attack.

Password usage

Fact is that in spite of all the great crypto around, a significant fraction of our security today resides in passwords: bank ATM passwords; login passwords; passwords for different websites; ...

Few of us today have cryptographic keys; but we all have more passwords than we can remember!

Passwords are convenient and entrenched.

Preventing dictionary attacks is an important concern.

Preventing dictionary attacks: Password selection

Systems try to force users to select "good" passwords, meaning ones not in the dictionary. But studies show that a significant fraction of user passwords end up being in the dictionary anyway.

Attackers get better and better at building dictionaries.

Good password selection helps, but it is unrealistic to think that even the bulk of passwords are well selected, meaning not in the dictionary.

Preventing dictionary attacks: avoiding image revelation

An alternative approach is to ensure that usage of a password pw never reveals an image $\overline{pw} = f(pw)$ of pw under a public function f. Then, even if the password is in the dictionary, the dictionary attack cannot be mounted.

Password-based session-key exchange:

- A, B share a password pw.
- They want to interact to get a common session key.
- The protocol should resist dictionary attack: adversary should be unable to obtain an image of pw under a public function.

Doesn't prevent dictionary attacks

$$
\begin{array}{ccc}
A^{pw} & & B^{pw} \\
 & \xrightarrow{\quad A,\, g^a \quad} & \\
 & \xleftarrow{\quad B,\, g^b,\, \overbrace{\mathrm{MAC}_{pw}(1, A, B, g^a, g^b)}^{\sigma} \quad} & \\
 & \xrightarrow{\quad A,\, \mathrm{MAC}_{pw}(0, A, B, g^a, g^b) \quad} &
\end{array}
$$

Session key is $K = H(A, B, g^a, g^b, g^{ab})$.

Dictionary attack is possible: Let f be defined by

$$f(x) = MAC_x(1, A, B, g^a, g^b)$$

Then, letting D be the dictionary of candidate passwords, get the target password pw via

for all $pw' \in D$ do
    if $f(pw') = \sigma$ then return pw'

Protocol KE4: [BPR00]

$$
\begin{array}{ccc}
A^{pw} & & B^{pw} \\
 & \xrightarrow{\quad A,\, E_{pw}(g^a) \quad} & \\
 & \xleftarrow{\quad B,\, E_{pw}(g^b),\, H(1, A, B, g^a, g^b, g^{ab}) \quad} & \\
 & \xrightarrow{\quad A,\, H(2, A, B, g^a, g^b, g^{ab}) \quad} &
\end{array}
$$

$E : PW \times G \to G$ is a block cipher over group G and keyspace PW of all possible passwords; the session key is $K = H(0, A, B, g^a, g^b, g^{ab})$.

This has been proven secure against dictionary attack [BPR00].