Email Security and Wireless Security

CS4264
Gang Wang
E-MAIL SECURITY

Slides credit to Cliff Zou
Email Security

- Security problem: Spoofing
- Security problem: Spam
- PGP (pretty good privacy)
Recap: SMTP

Three major components:
- User agents (client)
- Mail servers
- Simple mail transfer protocol: SMTP

Email Clients
- Composing, editing, reading mail messages
- Retrieve and send email through mail servers
Mail Server

- Some SMTP Commands:
  - MAIL FROM: <reverse-path>
  - RCPT TO: <forward-path>
  - RCPT TO: <forward-path>
    - If unknown recipient: response "550 Failure reply"
  - DATA
    - email headers and contents

- Use TCP port 25 for connections
Spoofed Email

• **SMTP**: designed for a trusting world

• Data in **MAIL FROM** totally under control of sender
  
  … an old example of improper input validation

• Recipient’s mail server:
  
  – Only sees **IP address** of direct peer
  
  – Recorded in the first **From** header
Example: Spoofed Email

support_ref_5581@ebay.com

http://theconversation.com/in-cybersecurity-the-weakest-link-is-you-33524
Behind that Email

Microsoft Mail Internet Headers Version 2.0

Received: from mer-w2003-6.napier-mail.napier.ac.uk ([146.176.223.1]) by EVS1.napier-mail.napier.ac.uk with Microsoft SMTPSVC(6.0.3790.1830);

Wed, 18 Jan 2006 00:17:45 +0000

Received: from pcp0011634462pcs.ivylnd01.pa.comcast.net (Not Verified[68.38.82.127]) by mer-w2003-6.napier-mail.napier.ac.uk with NetIQ MailMarshal (v6,1,3,15)

id

FCC: mailbox://support_id_1779124147875@ebay.com/Sent

Date: Tue, 17 Jan 2006 17:10:39 -0700

From: eBay
Demo

- [https://emkei.cz/](https://emkei.cz/)
- Send a spoofed email
What is Spam?

• Unsolicited bulk email messages

“Spam is e-mail that is both unsolicited by the recipient and sent in substantively identical form to many recipients”

• As of last quarter of 2005, estimates indicate that about 80-85% of all email is spam

• Microsoft founder Bill Gates receives four million e-mails per year, most of them being spam
Spamming Methods

- **Direct spamming**
  - By purchasing upstream connectivity from “spam-friendly ISPs”

- **Open relays and proxies**
  - Mail servers that allow unauthenticated Internet hosts to connect and relay mail through them

- **Botnets**
  - Collection of machines acting under one centralized controller. Eg: Bobax

- **BGP Spectrum Agility**
  - IP hijacking techniques
Migration and Anti-spam

• Filtering
  – Based on content
  – Use features in email’s headers and body
  – Eg: SpamAssassin

• Blacklisting
  – IP addresses of known spam sources
  – More than 30 widely used blacklists available today
Still Not Perfect Solution

• Content-based properties are *malleable*
  – **Low cost to evasion**: Spammers can easily alter features of an email’s content
  – **Customization**: Customized emails are easy to generate
  – **High cost to filter maintainers**: Filters must be continually updated as content-changing techniques become more sophisticated

• Content-based filters are *applied at the destination*
  – **Too little, too late**: Wasted network bandwidth, storage, etc. Many users receive (and store) the same spam content

• Blacklists
  – Aggressive filters have many false positives
  – One list might not have all the information about spamming IPs
  – Need to consult multiple lists
Spam Detection by Network Traffic?

Network-level properties of spam arrival
• From where?
  – What IP address space?
  – ASes?
  – What OSes?

• What techniques?
  – Botnets
  – Short-lived route announcements
  – Shady ISPs

• Capabilities and limitations?
  – Bandwidth
  – Size of botnet army
So much can be done by email service providers …

How to do end-to-end security for our own emails?
PGP: Pretty Good Privacy

- Functionality
  - Encryption for confidentiality
  - Signature for non-repudiation/authenticity
- Sign before encrypt, so signatures on unencrypted data - can be detached and stored separately.
- PGP-processed data is base64 encoded

- First released in 1991,
- developed by Phil Zimmerman
- Free software: OpenPGP and variants
- OpenPGP specified in RFC 2440
PGP Services

PGP supplies 5 basic services:
1. Authentication
2. Confidentiality
3. Compression
4. Email compatibility
5. Segmentation

Actually, only authentication and confidentiality are really “services”. The others are engineering features designed to make PGP efficient and robust.
PGP Authentication

This is a digital signature function.

1. Sender creates a message $M$.
2. Sender generates a hash of $M$.
3. Sender signs the hash using his private key and prepends the result to the message.
4. Receiver uses the sender’s public key to verify the signature and recover the hash code.
5. Receiver generates a new hash code for $M$ and compares it with the decrypted hash code.

Abstractly:

$$S \rightarrow R : \{h(M)\}_{K_s^{-1}}, M$$
PGP Confidentiality

PGP provides encryption for messages sent or stored as files

1. Sender generates a message $M$ and a random session key $K$.
2. $M$ is encrypted using key $K$.
3. $K$ is encrypted using the recipient’s public key, and prepended to the message.
4. Receiver uses his private key to recover the session key.
5. The session key is used to decrypt the message.

Abstractly:

$$S \rightarrow R : \{K\}_{K_r}, \{M\}_K$$

But why? Why not just use public key to encrypt the message instead of session key?
Confidentiality and Authentication

Both authentication and confidentiality may be combined for a given message.

1. Apply the authentication step to the original message.
2. Apply the confidentiality step to the resulting message.

• Should we sign-then-encrypt or encrypt-then-sign?
Key Management

PGP makes use of four types of keys: one-time session symmetric keys, public keys, private keys, passphrase-based symmetric keys.

Session keys: used once and generated for each new message
Public keys: used in asymmetric encryption
Private keys: also used in asymmetric encryption
Passphrase-based keys: used to protect private keys

A single user can have multiple public/private key pairs.
Why So Many Keys?

- PGP uses four kinds of keys: session keys, public and private keys, and passphrase generated keys.

- Public / private key pairs are the most expensive to generate.

- Since the security of the system depends on protecting private keys, these are encrypted using a passphrase system.

- In PGP, session keys and passphrase-based keys are generated on the fly, used once and discarded.
One more thing about Keys

- Each PGP user must manage his own private keys and the public keys of others.

- These are stored on separate keys rings.

- Private keys are protected by encryption; public keys are stored with certificates attesting to their trustworthiness.

- Keys can be revoked.
WIRELESS SECURITY

Slides credit to Kurose & Ross
Wireless Security

- WEP security vulnerability
- 802.11i improved security
IEEE 802.11 wireless LANs (Wi-Fi)

Network infrastructure

wireless hosts
- Laptop, phone tablet
- Run applications
- Stationary or mobile
- Wireless does not always mean mobility

Link layer on the network stack
Same layer as Ethernet
Elements in Wireless Network

- **Ad-hoc mode**
  - no base stations
  - nodes can only transmit to other nodes within link coverage
  - nodes organize themselves into a network: route among themselves

- **Applications:**
  - battlefield networks,
  - networks in undeveloped countries
  - (one laptop per child OLPC)
  - Vehicular networks
One Laptop Per Child (OLPC)

Ad hoc networking among laptops with a few satellite connections

http://one.laptop.org/
802.11 Frame: Addressing

Internet

H1

AP

R1 router

AP MAC addr | H1 MAC addr | R1 MAC addr

address 1 | address 2 | address 3

802.3 frame

802.11 frame
IEEE 802.11 security

- **War driving**: drive around, see what 802.11 networks available?
  - More than 9000 accessible from public roadways, 85% use no encryption/authentication
  - Packet-sniffing and various attacks easy!
  - TJX data breach started from war driving (read P&P for more info)

- **Securing 802.11**
  - encryption, authentication
  - first attempt at 802.11 security
  - *Wired Equivalent Privacy (WEP)*: a failure

- **Current attempt**: 802.11i (WAP)
WEP data encryption

WEP is based on RC4 (Rivest Cipher 4, designed in ’87)

RC4 is a stream cipher

- Host/AP share 40 bit symmetric key (semi-permanent)
- Append 24-bit initialization vector (IV) to create 64-bit key
- 64 bit key used to generate stream of keys, $k_i^{IV}$
- $k_i^{IV}$ used to encrypt the i-th byte, $d_i$, in frame:
  - $c_i = d_i \oplus k_i^{IV}$
- IV and encrypted bytes $\{c_i\}$ sent in frame
802.11 WEP encryption

- Stream cipher with symmetric keys

IV (per frame)
K_S: 40-bit secret symmetric
plaintext frame data plus CRC

key sequence generator (for given K_S, IV)

\( k_1^{IV} \) \( k_2^{IV} \) \( k_3^{IV} \) \( \ldots \) \( k_N^{IV} \) \( k_{N+1}^{IV} \) \( \ldots \) \( k_{N+1}^{IV} \)

\( d_1 \) \( d_2 \) \( d_3 \) \( \ldots \) \( d_N \) \( CRC_1 \) \( \ldots \) \( CRC_4 \)

\( c_1 \) \( c_2 \) \( c_3 \) \( \ldots \) \( c_N \) \( c_{N+1} \) \( \ldots \) \( c_{N+4} \)

Sender-side WEP encryption

IV and encrypted bytes sent in frame
CRC: cyclic redundant code for error detection
Risk of Re-used Keys in Stream Cipher

- Ciphertext1 XOR Ciphertext2 = plaintext1 XOR plaintext2
- Plaintext can be recovered through statistical analysis
Fundamental problem: $k_i^{IV}$ should never be reused

- WEP is based on RC4 that is secure if keys are used just once
- 24 bit IV: $2^{24} = 16,777,216$ possible combinations
- However, because of birthday paradox, we know collision with $\geq 50\%$ after seeing about $2^{12} = 4096$ IVs – a coarse estimation

- Breaking 104 bit WEP in less than 60 seconds
  - Recovered a 104-bit WEP key using less than 40,000 frames, success probability of 50%
  - To succeed in 95% of all cases, 85,000 packets are needed.
  - Packets sniffing takes one minutes with ARP reinjection
Vulnerable Integrity Verification Checking (ICV) in WEP

\[ \text{crc}(x \oplus y \oplus z) = \text{crc}(x) \oplus \text{crc}(y) \oplus \text{crc}(z) \]

Cyclic redundancy code (CRC)

- CRC is designed for error detection (0→1?)
- Not secure against intentional manipulations
- CRC-32 ICV is a linear function.
- Attacker can manipulate the stream cipher encrypted CRC without the knowledge of the key
Summary of WEP Weaknesses

- **Weakness: Key Management and Key Size (40-bit)**
  - IEEE standard written in 1997, underestimated attackers’ power

- **Weakness: The Initialization Vector (IV) is Too Small (24-bit)**
  - Choosing IV randomly: (B-day paradox) 50% chance of reuse after less than 5000 frames
  - Choosing IV sequentially: when IV is reused, the keys are the same (predictable)

- **Weakness: Integrity Check Value is not appropriate (CRC-32)**
  - CRC-32 is for detecting inadvertent transmission errors (not for malicious attack models)
  - Forging ICV on tampered messages is possible
Improving Security in 802.11i

- Numerous (stronger) forms of encryption
- Provides key distribution/management
- Uses authentication server separate from access point
  - E.g., authentication server hosted by an enterprise
Wi-Fi Protected Access (WPA) & TKIP

• 128 bit key and 48 bit initialization vector

• Still uses RC4
  – For backward compatibility purposes, re-use existing hardware implementation

• Uses TKIP (temporal key integrity protocol)
  – TKIP implements a sequence number for each frame to prevent replay attacks
  – A key derivation function to generate per-frame key

• A stronger Message integrity checking (MIC) than CRC32
802.11i: Four Phases of Operations

**STA:** client station

**AP:** access point

1. Discovery/negotiation of security capabilities

2. STA and AP mutually authenticate.

3. STA derives Pairwise Master Key (PMK)

3. AP derives same PMK

4. STA, AP use PMK to derive Temporal Key (TK) used for message encryption, integrity

Data transmission with TK ……
802.11i WPA2

• WPA2 is an improvement over WPA
  – No longer back-ward compatible, requires new hardware!
• WPA2 uses AES (a block cipher)
  – instead of RC4 (a stream cipher) + TKIP
• AES encryption not compatible with legacy devices;
  – 128, 192, or 256 bit keys
  – Either needs new chipset to work,
  – or introduce an authentication server
• Same authentication mechanism

Dictionary attacks are possible against WPA and WPA2
• master key is computed from passphrase and SSID
• session keys from MAC addresses, nonce