# Information and Data literacy

[CB41216] Thurs 8:50-10:20

# PART I: Introduction 2. Basics usage of essential information systems

2<sup>nd</sup> part: network & Cybersecurity

Xavier Dahan , dahan.xavier.a2@tohoku.ac.jp
Some slides were made by Shuji ISOBE and Eisuke KOIZUMI from Center for Data Driven
Science and AI

#### (1) Common Tool

- Common Tools : Web Browser
- Common Tools : Internet
- URL
- HTTP
- HTTPS = HTTP + SSL/TLS
- Home Page and Website
- Common Tools : Email
- Using and Writing emails
- SPAM
- Common Tools : Text Editor
- Common Tools : Office Suite
- Common Tools : Console

#### (2) Filesystem

- Files
- Directory
- Filesystem

- Path
- Current Directory
- Home directory
- Change directory
- Directory tree of the ICL lab

#### (3) Network

- Internet and Protocol
- Protocols of transmission
- Examples of protocols
- IP addresses
- Domain and DNS
- Default subnet mask
- Ifconfig (Linux) Ipconfig(Windows)
- Routing, traceroute/tracert

#### (4) CyberSecurity

- Where security matters
- Security for Web Services

#### • Threat 1: Phishing

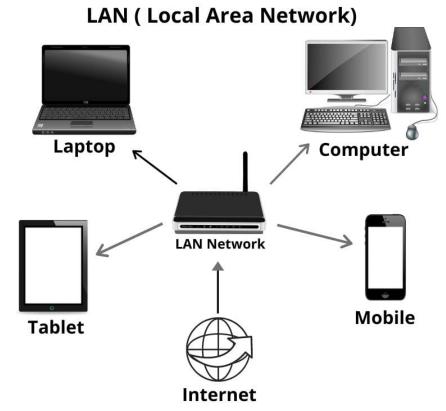
- countermeasure to Phishing 1:Domain Name
- Beware of URL "Homograph" attack
- countermeasure: Checking the protocol
- Countermeasure: Encryption and authentication
- Digital authentication: certificate
- Authentication on the user side
- Checking certificates with the browser
- Encryption + Authentication = safe ? Not always!
- Authentication: Reputation of a certificate
- Threat 2: Targeted Mail Attack
- Countermeasure
- Common Pattern
- "Ransom" Attack
- Threat 3: malware
- Malware: How do we be come infected?
- Countermeasure to Malware Infection
- Threat 4: Intrusion

### Internet and Protocol

- Internet: refers to the interconnection of all networks based on a set of common protocols agreed worldwide.
- Internet allows to unify all networks into one unique huge network.

### Local Area Network (LAN):

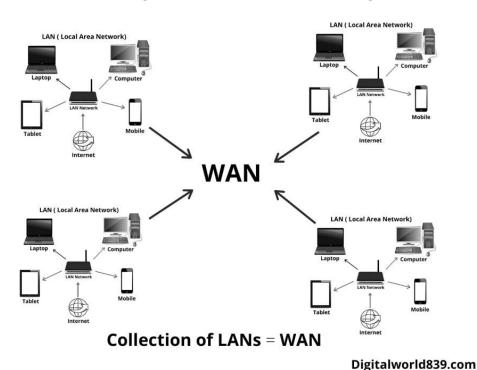
- Devices connected together inside a same place (small building or school, ..)
- Communications outside the LAN are set by the sysadmin (system administrator).

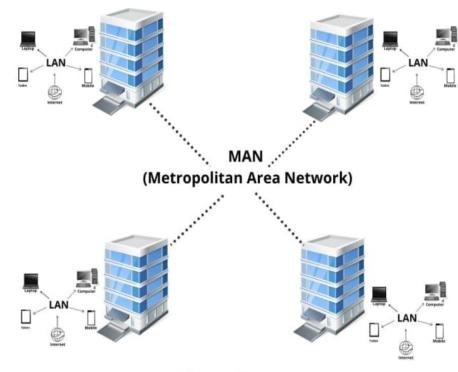


### Metropolitan Area Network (MAN)

- MAN connects some related LAN subnetworks into a same big network
  - ex: all departments of a company and its local branches
  - ex: departments of a large university

#### **WAN (Wide Area Network)**





Digitalworld839.com

### Wide Area Network (WAN):

WAN connects a collection of LAN networks over a wide area (ex: Internet)

### Protocols of transmission

#### **Protocol:**

- set of rules, agreed in advance (standardised) between all parties, to send and receive data across a network.
- A LAN network can connect to a WAN network (and thus internet) through such protocols.
- The protocols include the kind of cables (ethernet cable, optical etc. ) and the devices (switch, router) to be used across the networks traversed.

#### Packet:

- format of the data sent across the network that contains the message to be sent.
- The message is encapsulated by several layers that comply with the various protocols and devices crossed when navigating through the web.

### Examples of protocols

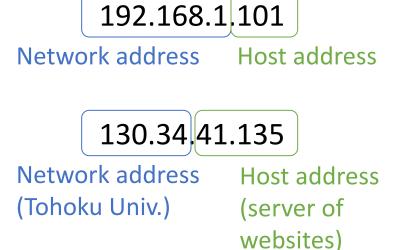
TCP/IP: set of protocols for data transmission over the internet. Examples of protocols contained in TCP/IP:

- SMTP (Simple Mail Transfer Protocol) : send emails
- POP (Post Office Protocol): receive emails
- IMAP (Instant Message Access Protocol): receive emails
- HTTP (HyperText Transfer Protocol): webpage's data transmission
- FTP (File Transfer protocol)

• ....

### IP addresses

- Sequence of four 8bits=1byte numbers (between 0 and 255).
- IP addresses have two parts:
  - Network address (first part, or suffix). Address of a LAN.
  - Host address (last part, or prefix). Address of a device inside the LAN.
- Subnet mask:
  - Determines how many bits of the IP address make the Network address.
  - Common: 255.255.255.0 The network can host around 250 IP addresses (network address is then the first 24 bits)
  - Also common: 255.255.0.0
     The network can host up to 250\*250 IP addresses (network address is then the first 16 bits)



### Domain and DNS

- Type <a href="http://130.34.41.135/">http://130.34.41.135/</a> in the address bar of your browser.
- It directs to the homepage of Tohoku University



- IP address is not convenient to remember for human.
- www. denotes host name (directly followed by the domain name)

### Default subnet mask

- When you typed http://130.34.41.135/ in the address bar, no subnet mask was given.
- How did the browser recognize the network address part from the host address part inside that IP address?

If no subnet mask is provided, the first 8bit number determines a default subnet mask according to the following rule:

	first 8bits number xxxx	Default mask	Network address	Host address
Class A	$0 \le xxxx \le 127$	255.0.0.0	First 8 bits	Last 24bits
Class B	$128 \le xxxx \le 191$	255.255.0.0	First 16bits	Last 16bits
Class C	$192 \le xxxx \le 223$	255.255.255.0	First 24bits	Last 8 bits
D, E	Not used much			

# Ifconfig (Linux) Ipconfig (Windows)

**Linux** console: ifconfig IP address is 10.34.49.193 Subnet mask: 255.255.255.0

Network address: 10.34.49

Host address: 193

Windows console: ipconfig IP address is 192.168.0.109 Subnet mask: 255.255.255.0

Network address: 192.168.0

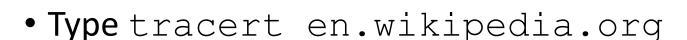
Host address: 109

### Router, route table

- The traffic on the network is essentially managed by routers.
- A router builds (and save) route tables.
- It stores all the IP addresses of routers traversed to reach the destination.
- The general principle of the construction route table is called a routing scheme.
- Over the internet, unicast is the most widespread routing scheme
  - Unicast: host/client (one node of the network to another node).
- A **routing protocol** (within the routine scheme) determines the algorithm that decides how the route table is going to be build

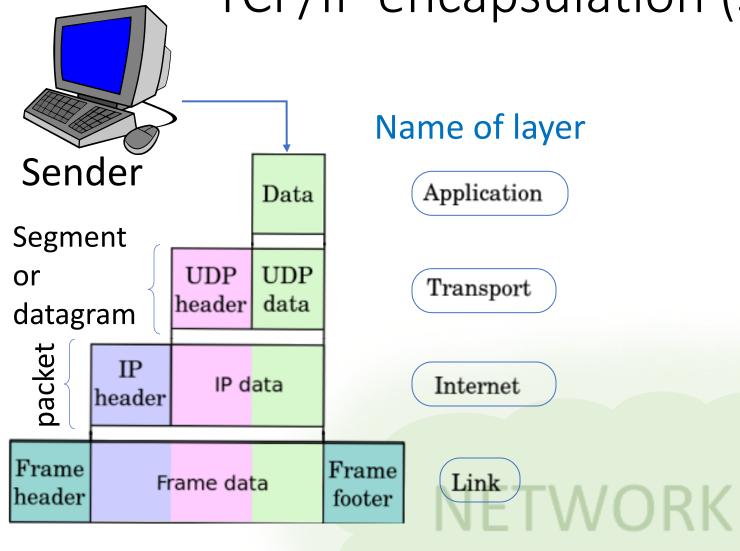
### traceroute (Linux) & tracert (Windows)

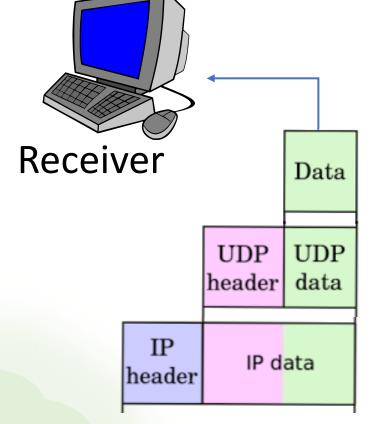
• Open a terminal (Windows key + cmd)



• To some extent (the security of Tohoku's university prevents to show the answers of some routers traversed), it shows the route to reach the English pages of Wikipedia.

 Gateway corresponds to the address to get out of the LAN you are in. TCP/IP encapsulation (simplified)





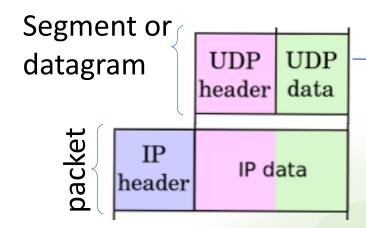
# Simplified and incomplete overview of layers

- Layer 5, Application: contains the data to be sent by the user, and by which protocol (Mail=SMTP or POP or IMAP?, HTTP, FTP etc.)
- Layer 4, Transport: adds a header to the message that contains flow control and system information (process-to-process delivery)-> segments (if UDP) or datagram (if TCP).
- Layer 3, Network: add a header that contains routing information across as many systems as necessary (hop-by-hop delivery) -> packets
   Layer 2, Datalink: add a header that contains information access to a neighbor system-> frame
- (Layer 1, Physical: convert to frame to 0101011... and it to the cables of the network)

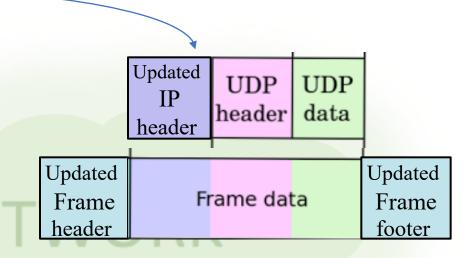
# TCP/IP hop-by-hop (simplified)

Intermediate
System (router)





- The transport layer is not touched
- Some intermediate system (switch)
  may only touch the frame at the
  datalink layer, not the packet



From previous system traversed

To next system traversed

#### (1) Common Tool

- Common Tools : Web Browser
- Common Tools : Internet
- URL
- HTTP
- HTTPS = HTTP + SSL/TLS
- Home Page and Website
- Common Tools : Email
- Using and Writing emails
- SPAM
- Common Tools : Text Editor
- Common Tools : Office Suite
- Common Tools : Console

#### (2) Filesystem

- Files
- Directory
- Filesystem

- Path
- Current Directory
- Home directory
- Change directory
- Directory tree of the ICL lab

#### (3) Network

- Internet and Protocol
- Protocols of transmission
- Examples of protocols
- IP addresses
- Domain and DNS
- Default subnet mask
- Ifconfig (Linux) Ipconfig(Windows)
- Routing, traceroute/tracert

#### (4) CyberSecurity

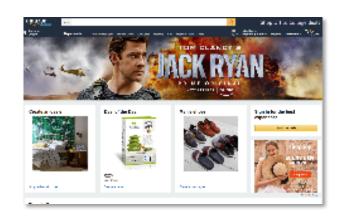
- Where security matters
- Security for Web Services

#### Threat 1: Phishing

- countermeasure to Phishing 1:Domain Name
- Beware of URL "Homograph" attack
- countermeasure: Checking the protocol
- Countermeasure: Encryption and authentication
- Digital authentication: certificate
- Authentication on the user side
- Checking certificates with the browser
- Encryption + Authentication = safe ? Not always !
- Authentication: Reputation of a certificate
- Threat 2: Targeted Mail Attack
- Countermeasure
- Common Pattern
- "Ransom" Attack
- Threat 3: malware
- Malware: How do we be come infected?
- Countermeasure to Malware Infection
- Threat 4: Intrusion

# Where security matters

There are many situations of security concerns in our daily "cyberspace life."



Online Shopping



Online Banking



**Online Communication** 



Campus Life



**Cloud Services** 

# Security for Web Services



**Online Shopping** 





Online banking

Any potential risk?

- These sites request you to send your secret information.
  - ★Name, address, phone number, card number, account number ....



Any communication over the Internet can be eavesdropped if it is not protected by some means (e.g. encryption.)

# Security for Web Services



**Online Shopping** 





Online banking

Any potential risk?

- These sites request you to send your secret information.
  - ★Name, address, phone number, card number, account number ....



What if the site is a "faked" one?

# Threat 1: Phishing attack

• Invite the "victims" to a fake website.

Let them input their personal/secret information, such as credit-card number,



## How does phishing work?

#### Why such an attack is possible?

- FACT: It is NOT hard (even for non-specialists) to build up a fake site which is indistinguishable from the genuine one on their appearance.
- If we weren't aware of that fact, we wouldn't expect in the first place that the website is fake.
  - ★ We wouldn't be aware of the potential risk.
- Concept of threats in the cyberspace is more abstract (virtual)
  - unaware users may act less carefully in the cyberspace.
    - ★ Are you always careful whenever you click on a link?

# Phishing: example

Existing Example: Shinsei Bank (Japan)





Which is the genuine one?

It is almost impossible to distinguish their appearance.

### countermeasure to Phishing 1: Domain Name

#### **Genuine URL of ISTU**

```
https://istu3g.dc.tohoku.ac.jp/istu3g/auth/login

Protocol Tohoku Univ's Domain Folder Names and Paths in the server
```

#### **URL** of a (Possibly) Fake Site

https://hogehoge.ne.jp/tohoku/istu3g/login

**NOT** Tohoku Univ's Domain

These are merely folder names. Anyone can claim "tohoku" and "istu3g" identities here.

# Beware of URL "Homograph" attack

- Checking the URL is a good start but...
- ...there are tricks for faking URLs.
- Several letters in the URL are replaced with different ones, but they look almost the same.
- It is hard to distinguish the faked URL from the genuine one.

```
www.apple.com
```

```
www.appie.com
www.apple.com
www.apple.com
```

 $\leftarrow$  hard to find a fake

### countermeasure to Phishing 2: Checking the protocol

Sensitive data shall be encrypted: HTTPS protocol see <u>this</u> slide)

#### **Key Features of HTTPS**

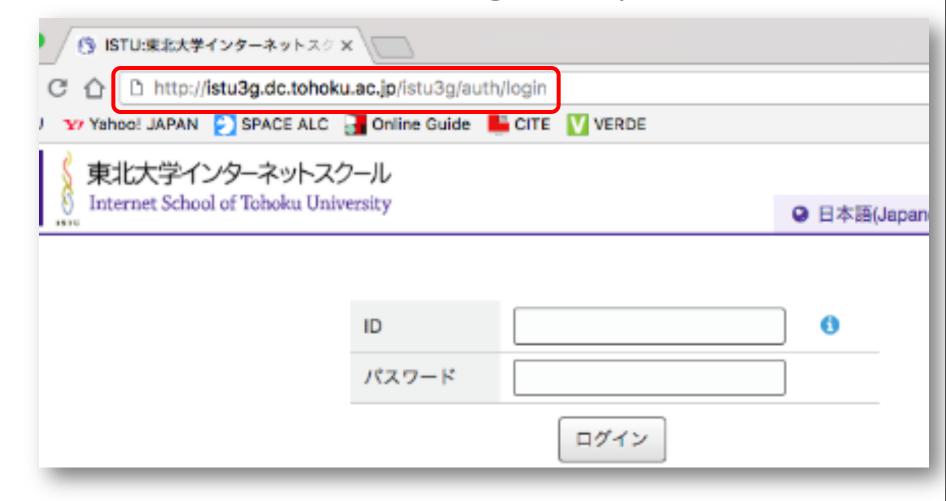
- *Encryption*: prevent eavesdropping
- Certificates Verification (digital authentication): verify the validity of the server's identification.

### Why "encryption" is needed?

 Any communication over the Internet is exposed to the risk of being eavesdropped if no countermeasure is implemented

### countermeasure to Phishing 2: Checking the protocol

If the website requests us to input secret information such as ID/PW but does not implement the HTTPS protocol, it may be a fake one, or simply insecure.



Is encryption sufficient? The answer is "NO."

### Phishing countermeasure 3: Encryption and authentication

#### Why is encryption not sufficient?

- The server you are now connecting to is **NOT** necessarily the genuine one.
- 2 Even malicious attackers can establish an HTTPS channel with you.

#### The second countermeasure: authenticate the host of the webpage

Besides encryption, HTTPS implements a layer of digital authentication.

If the authentication succeeds, one can be convinced that it is the "genuine" server with high (but not total) confidence.

### Principle of digital authentication: certificate

The server hosting the website asks for a **certificate** to an authority:



1 Request to issue certificate



Certificate Authority (CA)





4 Issue the certificate

- ② Examine the request
- 3 Generate the certificate

5 Install the certificate on the server.



### Principle of digital authentication: Certificate issuance

#### **Certificate**



- Issued by a Certificate Authority (CA), carries its signature
  - The CA generates this digital signature by using its "secret signing key."
  - This digital signature is **uniquely attributed** to the document (here a website) to be signed.
  - Without this secret key, it is **infeasible** to forge a valid signature.

**Crucial Premise** ①: *Only the CA can generate a valid signature.* 

**Crucial Premise** ②: The operation of authentication can be done by anyone with the "verification public key" uniquely associated to the secret signing key.

These premises can, to some great extent, be realized by "Public-Key Cryptography."

### Authentication on the user side



1 Request for connection



Website

3 Verify the signature of the certificate

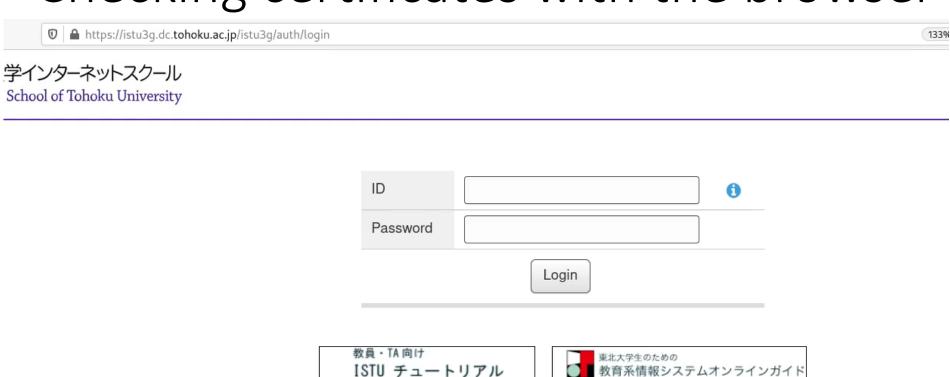
2 Present the attributed certificate



Signed with the CA's Secret signing key

Use the "public verification key", which is built in major web browsers

## Checking certificates with the browser



本システムは、8月末で停止予定です。新しいISTUについては、こちらをご覧下さい。

Online Guide: Systems & Services for Students in Tohoku Universit

オンライン授業の情報をこちらにまとめましたのでご覧ください。

ISTU tutrial for teacher and teaching assistant

### Encryption + Authentication = safe ? Not always

**Problem:** What if the CA is not trustworthy?

**FACT:** Establishing a CA can be done by anyone, even by the malicious attackers (if they make some effort.)



A "valid" certificate is insufficient, because it may be issued by untrusted (or even malicious) CA or be "self-signed."

It is risky to accept such "self-signed signatures."

### Authentication Reputation of a certificate

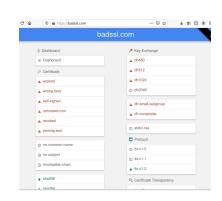
**Problem:** What if the CA is not trustworthy?

#### A Simple Countermeasure:

We accept those CAs which have been socially acknowledged as trustworthy authorities: builds a reputation.

The current major browsers show warning messages when they detect certificates issued by *untrusted* or *unknown* CAs.

**Example:** badssl.com hosts (on purpose) sites with invalid CA. Let's examine how firefox reacts when a connection is attempted.



### Threats on the network 2:







Defining the target e.g. students in Tohoku Univ.

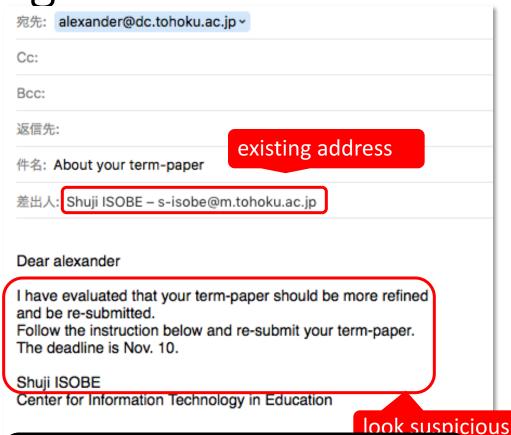


- (2) Send "invitation" mails. disguising authentic ones "Oh, it's from my teacher."
- 3 Click on the URL (to the hell) and send secret data



4 Attack completed

Your secret has been handed to the attacker!



Instruction for re-submission:

- Visit the following website: http://www.math.tohoku-u.ac.jp/lsobe/info-B/submission-form.html
- 2. Login with your ID and PW.

3. Follow the instruction on the sc Invitation to the hell

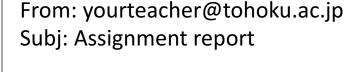
# Targeted Mail Attack (with attachment)

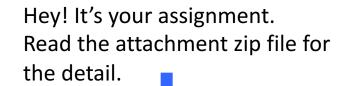


Defining the target e.g. students in Tohoku Univ.



② Send "invitation" mails.







Open the attachment.



Malicious program infection (but hard to be aware of)



Attack completes and goes on





# Targeted Mail Attack

#### Is it a kind of "SPAM"?

The targeted attack is not the same as the simple spams.



- Target the "victim" specifically
- Persistent, specific purpose
- Advanced, and carefully designed and prepared tricks





- Tend to accept "invitations" more easily.
  - ★They disguise themselves as trustworthy messages.
- Hard to detect (even by "antivirus")

The targeted attack is not like indiscriminate SPAM. They have specific purpose and target.

## Countermeasure 1

## Point 1 Is "FROM" trustworthy?

The "FROM" header can be forged.



Even if "FROM" claims "tohoku.ac.jp" domain, it might come from any other domain.

The sender's account might have been hijacked



The "real" attacker scatters the invitations from the "hijacked" account.

"FROM" is not necessarily trustworthy.

宛先: alexander@dc.tohoku.ac.jp ~		
Cc:		
Bcc:		
返信先:		
件名: About your term-paper existing address		
差出人: Shuji ISOBE – s-isobe@m.tohoku.ac.jp		
Dear alexander  I have evaluated that your term-paper should be more refined and be re-submitted.  Follow the instruction below and re-submit your term-paper. The deadline is Nov. 10.  Shuji ISOBE Center for Information Technology in Education		
Instruction for re-submission:  1. Visit the following website: http://www.math.tohoku-u.ac.jp/lsobe/info-B/submission-form.html		
2. Login with your ID and PW.		
3. Follow the instruction on the screen.		

## Countermeasure 2

### **Point 2** Examine the contents

Is there anything unnatural?

- Language usage
- Message itself sounds strange
- suspicious URL (why not "tohoku.ac.jp" domain? Why not ISTU?)
- suspicious attachment (if exists)

夗元:	alexander@dc.tonoku.ac.jp *
Cc:	
Bcc:	
返信券	g:
件名:	About your term-paper
差出力	: Shuji ISOBE – s-isobe@m.tohoku.ac.jp

#### Dear alexander

I have evaluated that your term-paper should be more refined and be re-submitted.

Follow the instruction below and re-submit your term-paper. The deadline is Nov. 10.

Shuji ISOBE Center for Information Technology in Education

Instruction for re-submission:

- Visit the following website: <a href="http://www.math.tohoku-u.ac.jp/lsobe/info-B/submission-form.html">http://www.math.tohoku-u.ac.jp/lsobe/info-B/submission-form.html</a>
- 2. Login with your ID and PW.
- Follow the instruction on the screen.

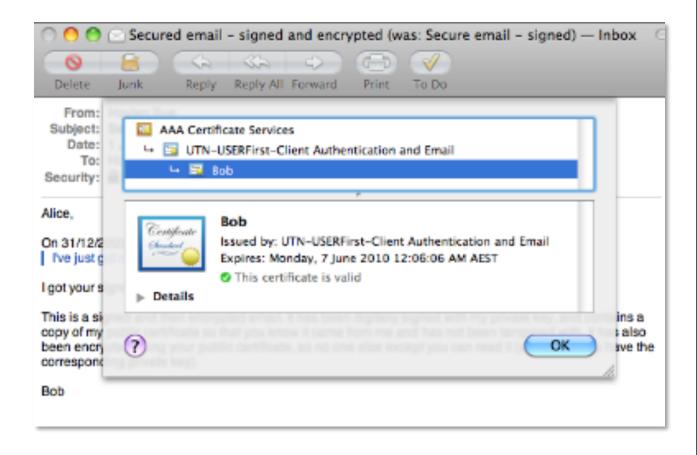
## Countermeasure 3

## **Point 3 Verify signatures**

Many socially-trusted organizations sends to the customers emails with signatures (certificates).



One can verify the signatures, as in the case of HTTPS communication.



#### (Repeat)

Be careful of "self-signed signatures" and "untrusted CA's signatures." If no signature is attached, it may be worth suspecting it.

# Common pattern of Targeted Mail Attack

## **Typical "Invitations"**

- Your SNS accounts (or other accounts, such as net bank, Tohokudai-ID, Apple ID etc.) were hijacked (or locked). Please visit the following website and update your password.
- A suspicious behavior has been detected on your account. Please visit ...
- We improve the security of the online banking system. Visit the following link ...
- Our service suffered from information leakage incident lately. Please visit ...

### Be careful, especially if the "invitation" includes

- request for visiting unknown or suspicious hyperlinks
- request for typing passwords or secret codes

## "Ransom" Attack

#### **Overview of the attack**

- 1 Your PC is infected with a malware.
- 2 The malware "locks" your PC, for example, by locking the screen or encrypting the folders of your computer.
  - You can no longer operate your PC unless it is unlocked.
- ③The malware (attacker) requires you to pay a "ransom" in order to "unlock" your PC.



A ransomware "WannaCry" is now scaring you!

# Threats on the network 3: malware

### Malware (malicious software)

- Causes damage to data (destruction or corruption) and systems
- Gains unauthorized access to network or secrete data
- **Virus:** piece of malicious code that attaches to a clean (often executable) code in the system. When a user executes the program, it damages core functionalities of the system etc.
- Worms: Starts from one machine, spreads over the network.
- **Spyware:** Program hiding in the background that spies the computer. Collect information about the user.
- **Trojans:** (reference to Greek soldiers hidden in a giant horse) hides inside or disguises itself as a legitimate software. Creates backdoors in the system to give other malware easy access to the system.
- Ransomware: blocks the network/system unless a ransom is paid.

## Malware: How do we become infected?

1 Visiting "infected" websites.

unfixed "security holes"



attack!







- tampering the website
- embedding the malware's code



Even "legitimate" sites may be affected.



A single browsing can cause infection.





## Malware: How do we become infected?

## 2 Downloading/Executing

- Visit some websites (e.g. invited by some email.)
- Click some link to download some file
- Open the file, or executing the program

### Infection!



- You are convinced that these programs are innocent, or even useful/necessary.
- •Those malicious programs may even disguise themselves as "security tools."

(fake security tools)



a "fake" security tool

## Malware: How do we become infected?

- **3** Other possible infection routes
  - Open the files attached to some email
    - automatic opening is risky
  - •Insert (infected) USB devices or other removable media, and copy files
    - off-line infection
  - File sharing services

    Even small number of infected files may cause severe damage.
  - Macros
    - e.g. Macros built in the office-suite files may behave as malware.

## Countermeasure to Malware Infection

- •(Repeat) There is *no* perfect countermeasure.
- But, there are some basic countermeasures to take.
- 1 DO NOT download or execute unexamined files.
- ② Before opening the attachment files of emails, ask yourself whether or not the files are trustworthy.
- 3 DO NOT leave the security holes unfixed.
  - apply the "security patch" programs released by the trusted vender.
- 4 Use security suites such as "antivirus" or software firewalls.
  - e.g. detection of various malwares, monitoring communication

## Threats 4: Intrusion

- Password or ID leakage (vulnerable password: dictionary attack)
- Breach in the network architecture.
  - Example: Scan ports attack. A port is an address of a service within a system
     (≠ IP address → address within the network).
     The attack scans all ports and find open ports.
    - Try to find breaches, weakens guarding services to open doors ....
- Vulnerability in the implementation of a security service:
  - Example: buffer overflow. The place where a program writes data when still in execution is called a buffer. It has limited size. If the overflow is not planned in the implementation, malware can take advantage of an overflow.

#### (1) Common Tool

- Common Tools : Web Browser
- Common Tools : Internet
- URL
- HTTP
- HTTPS = HTTP + SSL/TLS
- Home Page and Website
- Common Tools : Email
- Using and Writing emails
- SPAM
- Common Tools : Text Editor
- Common Tools : Office Suite
- Common Tools : Console

#### (2) Filesystem

- Files
- Directory
- Filesystem

- Path
- Current Directory
- Home directory
- Change directory
- · Directory tree of the ICL lab

#### (3) Network

- Internet and Protocol
- Protocols of transmission
- Examples of protocols
- IP addresses
- Domain and DNS
- Default subnet mask
- Ifconfig (Linux) Ipconfig(Windows)

#### (4) CyberSecurity

- Where security matters
- Security for Web Services

#### Threat 1: Phishing

- countermeasure to Phishing 1:Domain Name
- Beware of URL "Homograph" attack
- countermeasure: Checking the protocol
- Countermeasure: Encryption and authentication
- Digital authentication: certificate
- Authentication on the user side
- Checking certificates with the browser
- Encryption + Authentication = safe ? Not always !
- Authentication: Reputation of a certificate
- Threat 2: Targeted Mail Attack
- Countermeasure
- Common Pattern
- "Ransom" Attack
- Threat 3: malware
- Malware: How do we be come infected?
- Countermeasure to Malware Infection
- Threat 4: Intrusion

## Table of Contents

### Common Tools (3-19)

- Web Browser
- Internet
- URL (Uniform Resource Locator)
- HTTP
- HTTPS = HTTP + SSL/TLS
- Home Page and Website
- Email
- Using and Writing email
- SPAM
- Targeted Attacks mail

- Text Editor
- Office Suite
- Console

#### Filesystem (20-25)

- Files
- Directory
- Filesystem
- Path
- Current directory, home directory
- Directory tree of the ICL lab computers

#### **Network (26-37)**

- Internet and Protocol
- Protocols of transmission
- Examples of protocols
- IP addresses
- Domain and DNS
- Default subnet mask
- Ifconfig (Linux) Ipconfig (Windows)
- Threats on the network 1: malware
- Threats 2: Intrusion
- Threats 3: Phishing websites

# Targeted Attacks mail

- Unlike SPAM, some emails you receive may look legitimate regarding the institution to which belongs the email address.
- Example: you are student at Tohoku University.
   You receive an email that looks like "From Administration" or from another student and say "Important Matters".
   For some reasons, it asks for some personal data.
   Or it has a file attached, and asks you to open it for more details.

 These attacks are more clever than SPAM and can be particularly vicious.

# Threats 3: Phishing websites

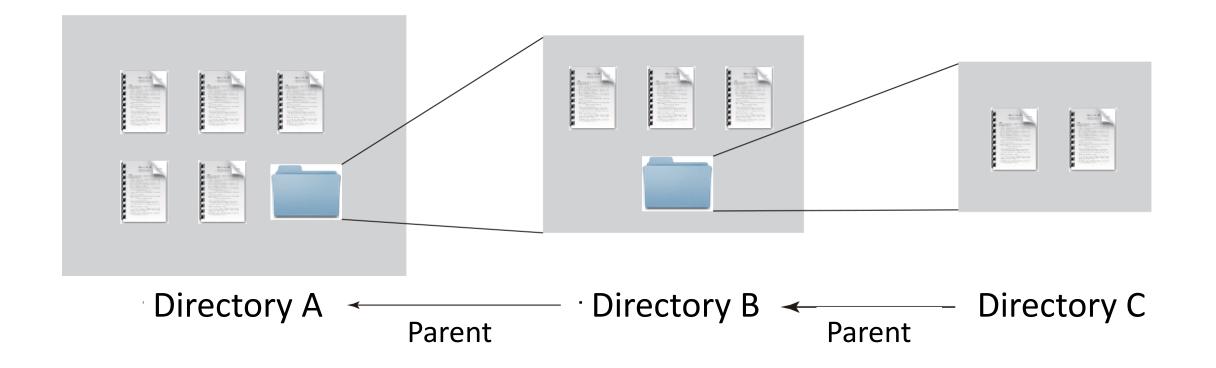
- False website that imitates famous websites (banks etc.)
- Users get confused and believe it is the authentic website.
   they may give personal information (credit card number etc.)
   these data are collected mostly for malicious purposes.
- An invitation to a phishing website is often sent though a "phishing email", which also imitates an email from a famous institution. This email contains links to a phishing website.
- (web) Scam: emails that try to connect you through human sensibility (ex: promises friendship from an attractive person ex: takes advantage of the loneliness of elderly persons etc.)

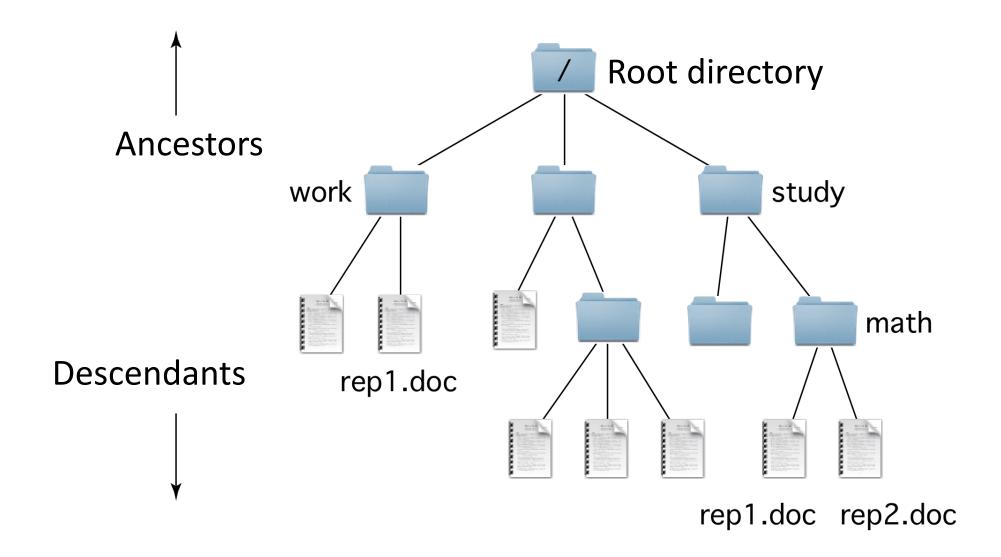
### How does malware spread?

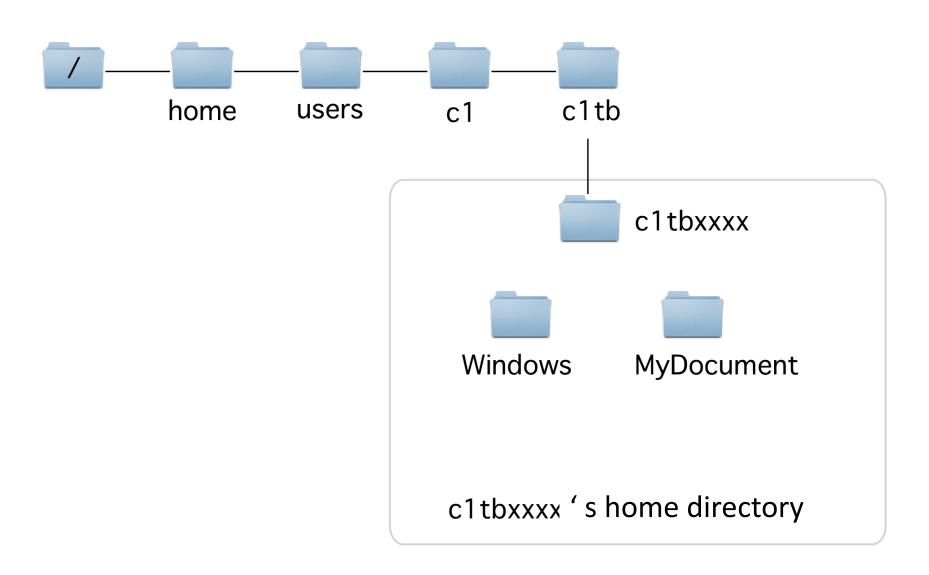
- Fraudulent email attachments. Or by clicking on a link in such an email.
- Delivered via instant messaging or social media
- Find a breach in a vulnerable software implementation. Etc.

Protection against malware.

- On the individual level:
  - First of all, be aware of threats.
  - Beware of weird emails, untrusted invitations on social media etc.
  - Perform regular updates of software (correction of breaches)
- Computer level
  - Install an antivirus (target not only viruses. Contains a catalogue of reported malware for inspection, and eventually destruction).
  - Necessary but only protect against known and reported malware.







# File Manager

• GUI to manipulate files (Windows: Explorer).