Anonymity on the Internet

Cunsheng Ding

HKUST

Hong Kong

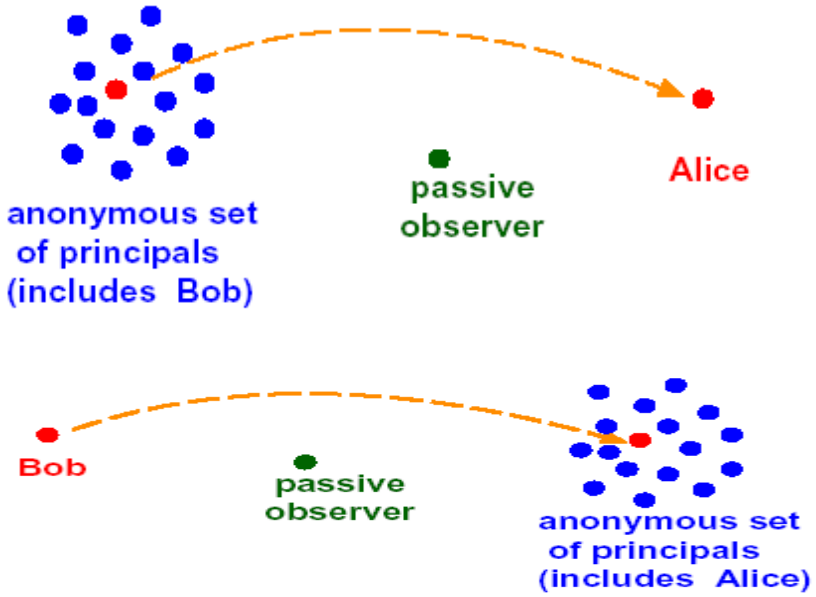
Part I: Introduction to Anonymity

Definition of Anonymity

- According to Wikipedia, anonymity is derived from the Greek word ἀνωνυμία, anonymia, meaning "without a name" or "namelessness".
- Anonymity of a subject means that the subject is not identifiable within a **set of subjects**, the anonymity set.
- The anonymity set may vary from time to time.

Types of Anonymity

- Sender anonymity
  - Receiver / observer can't identify sender
- Receiver anonymity
  - Observer can't identify receiver



The Demand of Anonymity: Before the Internet

- It is as old as identity and goes along with humankind from the dawn of history.
- Many acts of charity are performed anonymously, as benefactors do not wish, for whatever reason, to be acknowledged for their action.
- Journalists frequently use anonymous sources to help investigate news stories. As fear of the retaliation or persecution, the whistle-blowers do not want to disclose their identities. Then journalists report news using the anonymous sources, by quoting informants who were unwilling to reveal their identities as anonymous sources.

The Demand of Anonymity: The Internet Era

- The appeal of Internet communication depends **in part** on its capacity to support anonymity.
- Internet users can make political claims as well as non-political comments, engage in whistle-blowing, conduct commercial transactions, receive counseling and consume sexual materials without apparently disclosing their identities.
- Anonymous communication encourages Internet communications, and the Internet in turn may encourage anonymous communication.

The Negative Sides of Anonymity

- Spams
- DoS attacks (Denial of Service Attacks)
- Illegal activities
  - anonymous bribery,
  - copyright infringement,
  - harassment,
  - financial scams,
  - disclosure of trade secrets

Pseudonymity

- Definition of pseudonymity: Though the identity of the message sender may seem anonymous because it is not easy to uncover the sender's identity, but it is possible to discover the identity of the message sender.
- This kind of anonymity has significant benefits:
  - it enables citizens of a democracy to voice their opinions without fear of retaliation against their personal reputations, but it forces them to take ultimate responsibility for their actions should the need somehow arise.
- What we need is pseudonymity, rather than anonymity

Part II: Anonymity Mechanisms

Proxy

The first anonymity mechanism

Proxy (1)

- An **anonymous proxy** or **anonymizer** is a trusted middleman.
- Between source and destination, it acts as an intermediary and privacy shield to protect personal information by hiding the identity information of sender or receiver.
- It is used to protect the privacy and anonymity of web browsers from web site operators, Internet snoops, and even unfriendly governments.
- It makes activities on the Internet untraceable.
- It is the most commonly-used anonymity mechanism.

Proxy (2)

- Web client <-> Proxy Server <-> Web server
- The web server does not know the client, but the proxy server.
- The proxy server does know who the client is.
- A good web proxy will allow users to setup a ``TLS or SSL tunnel”. This will prevent packet sniffers from eavesdropping while the web client is surfing anonymously.

Proxy (3)

In addition to hiding IP address, an anonymous proxy server will typically remove traffic such as:

- Cookies
- Pop-ups
- Banners
- Scripts
- Referrer information

Mix

The second anonymity mechanism

Mix: Brief Introduction

- It was invented by David Chaum in 1981.
- It is the basic building block of nearly all modern high-latency anonymous communication systems.
- It is a process with which a Mix device
  - first accepts encrypted messages as input,
  - then decrypts them;
  - finally sends the decrypted version to the receiver

Mix: Technical Idea

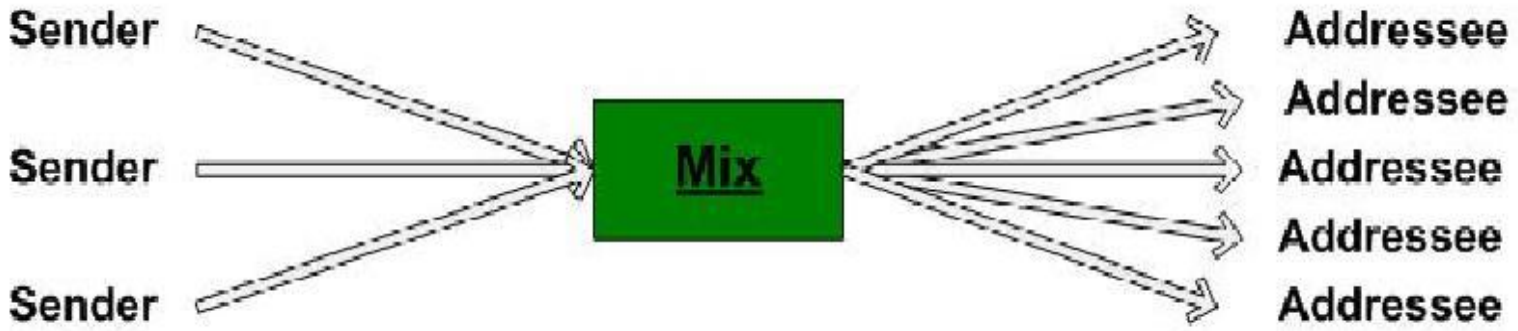
- Participant A prepares a message m for delivery to participant B by appending a random string R_0 to the message m, sealing it with the B's public key K_B, appending B's address Ad_B, and then sealing the result with the Mix's public key K_M. The Mix opens it with his private key, now he knows B's address Ad_B, and he sends $K_B(R_0,m)$ to B:

  $A \rightarrow K_M(R_1, K_B(R_0, m), Ad_B) \rightarrow Mix \rightarrow K_B(R_0, m) \rightarrow B$

- R_1 and R_0 are random strings which are used for preventing an attacker from guessing the message.

Mix: How to Use It

- The Mix will wait for a number of input messages, i.e., buffer them.
- The Mix then decrypts the messages with its private key.
- The decrypted messages will then be sent out in a different order



Mix: Security Questions

- Does it provide sender anonymity?
  - Yes. If the observer sees a message from the Mix to Bob, the observer does not know the sender's name. Similarly, Bob does not know the sender's name.
- Does it provide receiver anonymity?
  - Yes. If the observer sees a message from Alice to the Mix, the observer does not know the receiver's name.
- What will happen if only one sender is active for a window of time?

Differences between Proxy and Mix

- A text exchanged between a sender and a proxy may not be encrypted with a public key cipher, while a text exchanged between a sender and mix is encrypted with a public key cipher.
- A proxy will not wait for a number of messages to accumulate messages, while a mix will do so.

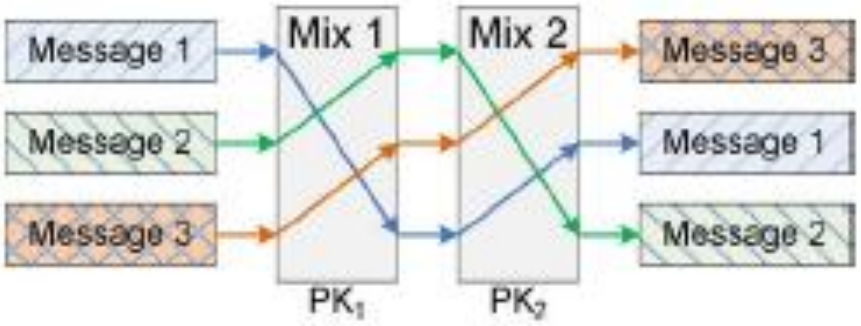
Mix Network

Fixed Topology (Cascade)

The third anonymity mechanism

Mix Network

- It is a network of mixes and used for providing better anonymity.
- The order of using the mixes is pre-determined and each mix is used exactly once.
- In case one of the mixes is not trustful, anonymity can still be achieved.



Mix Networks: Problems

- Fine for non-real-time applications (email)
- Not sufficient for VoIP (Voice over the Internet Protocol), video, Web
  - Mix waits to accumulate inputs to process as a batch (especially slow for low traffic)

Mix Networks: Enhancements

- Messages all the same length by padding
  - Why?
- Dummy messages inserted
  - Between mixes
  - Between mixes and user
  - Why?

Onion Routing

Dynamic Topology (Free-Route)

The fourth anonymity mechanism

Onion Routing (1)

- It is a distributed overlay network designed for anonymous communications over a computer network.
- It is based on the principle of Chaum's mix
- It encrypts message in layers, and there is an encrypted message for each mix on the path.
- Each message can take a different path through the network. The sender selects a random path.

Onion Routing (2)

- Each mix is a store-and-forward device that accepts fixed length messages, performs cryptographic operations on the messages, and then forwards the messages to the next node in the **routing path**. The mix is called an onion router.
- When an onion router receives a message, it knows the immediate predecessor of this message and to whom it should be passed.
- Each onion router removes a layer of encryption using its own private key, which will also uncover routing instructions of the next onion router.

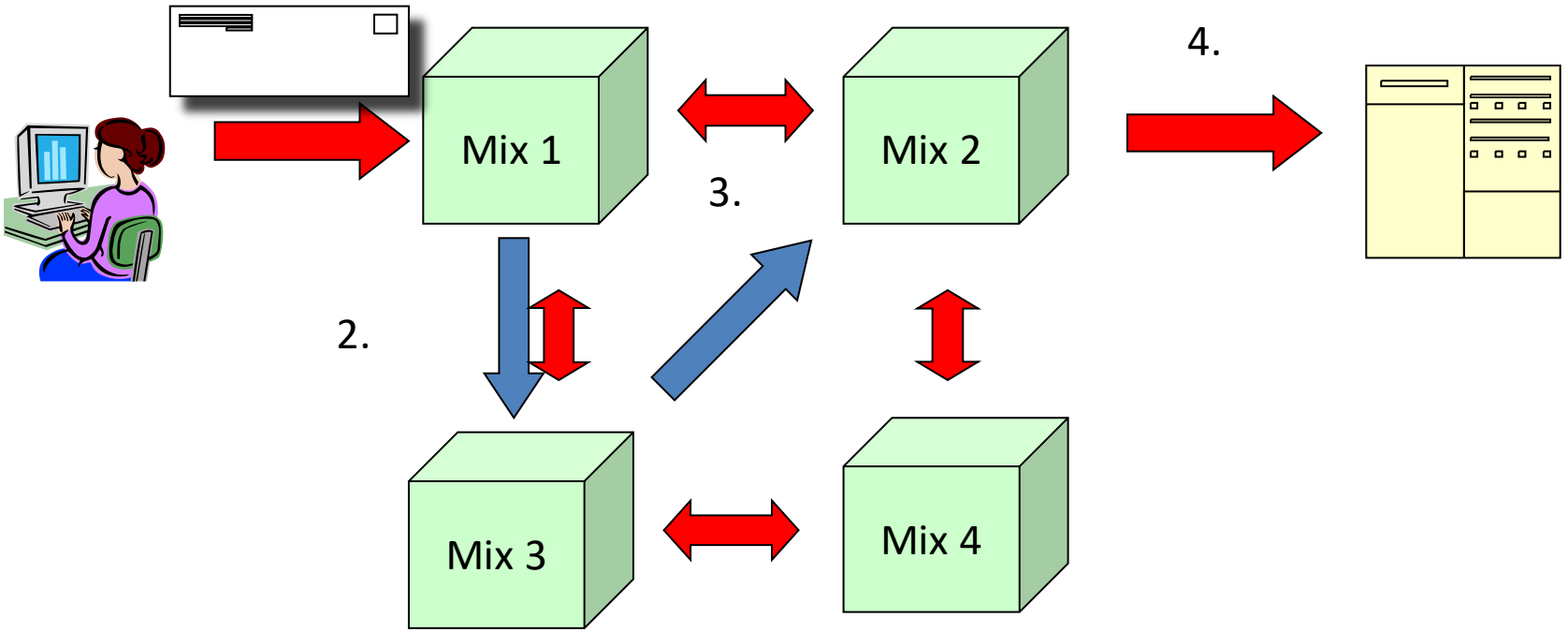
Ai = Next Hop Address

Ci = Message encrypted with public key of Mix i

S = Destination Host address

M = Original message

Each Mix is called an onion router



1. A1, **C1(A3, C3(A2, C2(S, M, r2), r3), r1)**
2. A3, **C3(A2, C2(S, M, r2), r3)**
3. A2, **C2(S, M, r2)**
4. S, **M**

The Mix Network versus the Onion Routing

- A Mix network uses several Mixes in a fixed order, and each Mix is used exactly once. A sender cannot choose the order of using the Mixes.
  - Hence, a text sent out from a sender or Mix has at most two layers of encryption.
- The Onion Routing uses a “multi-subset” of several Mixes and the order of using the multi-subset (i.e., the routing path) is selected by a sender.
  - Hence, a text sent out from a sender or Mix may have more than two layers of encryption.
  - A mix may be used twice or more times in a routing path.

Other Anonymity Mechanisms

- Tor: the most recent evolution of onion routing.
- Crowds: defense against internal receiver and a corrupted receiver, used for anonymous web browsing.
- DC-nets:
  - It is a secure multi-party computation protocol.
  - It provides sender and receiver anonymity without relying on a trusted third party.

Part III: Applications

Applications of Internet Anonymity

- Anonymous remailer
  - Anon.penet.fi, Cypherpunk, Mixmaster, Mixminion
- Anonymous electronic voting system
- Digital cash

Part IV: Regulations on Anonymity

EU Regulations

- A user should not be required to justify anonymous use.
- Anonymity should not be used as a cloak to protect criminals.
- Various countries have laws both protecting and forbidding anonymity.
- Different countries in EU have different laws.

Hong Kong Regulations

- Government agents are allowed to perform real identity trace in any communication channel including cyberspace if the anonymity is likely connected with crime or security measure.
- In general, persevering anonymity is recognized as a human right.
- Anonymity is not allowed for the sender for commercial purpose (no anonymous commercial emails).