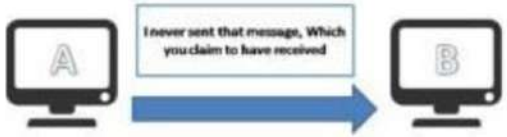
Electronic Mail Security

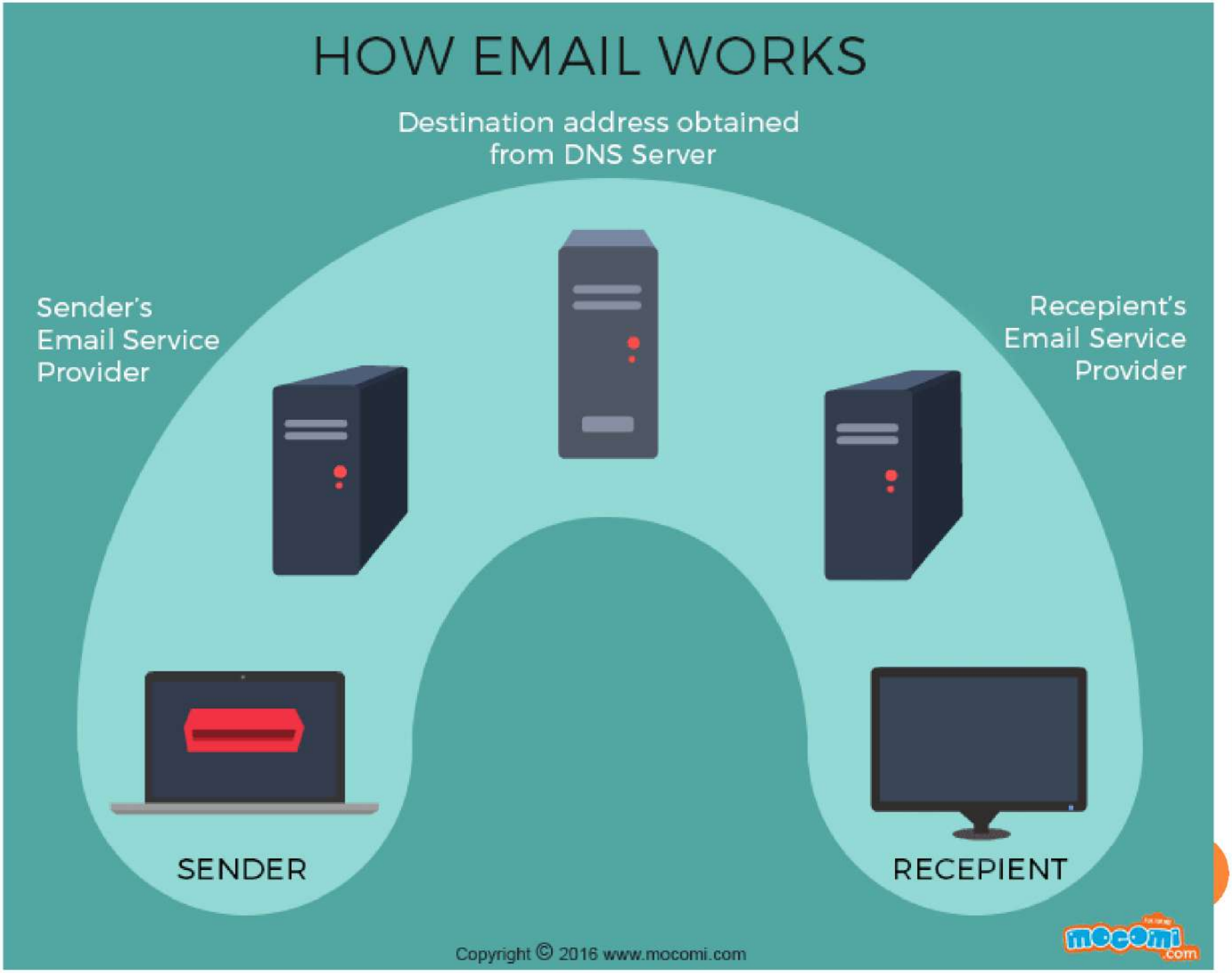
Ref: Cryptography and Network Security by William Stallings

Email Security

- Email is one of the most widely used and regarded network services
- Threats to the security of e-mail itself
  - Authentication
    - of sender of message
  - Non-repudiation of origin
    - protection from denial by sender
  - Loss of confidentiality
    - E-mails are sent in clear over open networks
    - E-mails stored on potentially insecure clients and mail servers
  - Loss of integrity
    - protection from modification



Non-Repudiation


HOW EMAIL WORKS
Destination address obtained from DNS Server
Sender's Email Service Provider
Recepient's Email Service Provider
SENDER
RECEPIENT
Copyright © 2016 www.mocomi.com

ABBREVIATION

- MUA – Message User Agents
- MHS – Message Handling Service
- MSA – Message Submission Agent
- MTA – Message Transfer Agent
- MDA – Message Delivery Agent
- MS – Message Store

INTERNET MAIL ARCHITECTURE

Sender

Transfer World

Receiver

MUA
Eg:
Outlook
Gmail

SMTP

MSA

SMTP

MTA

SMTP

Sender Mail Server
Eg: IBM server/ gmail server

Reciever Mail Server
Eg: Oracle server/ gmail server

MDA

Mail box

SMTP

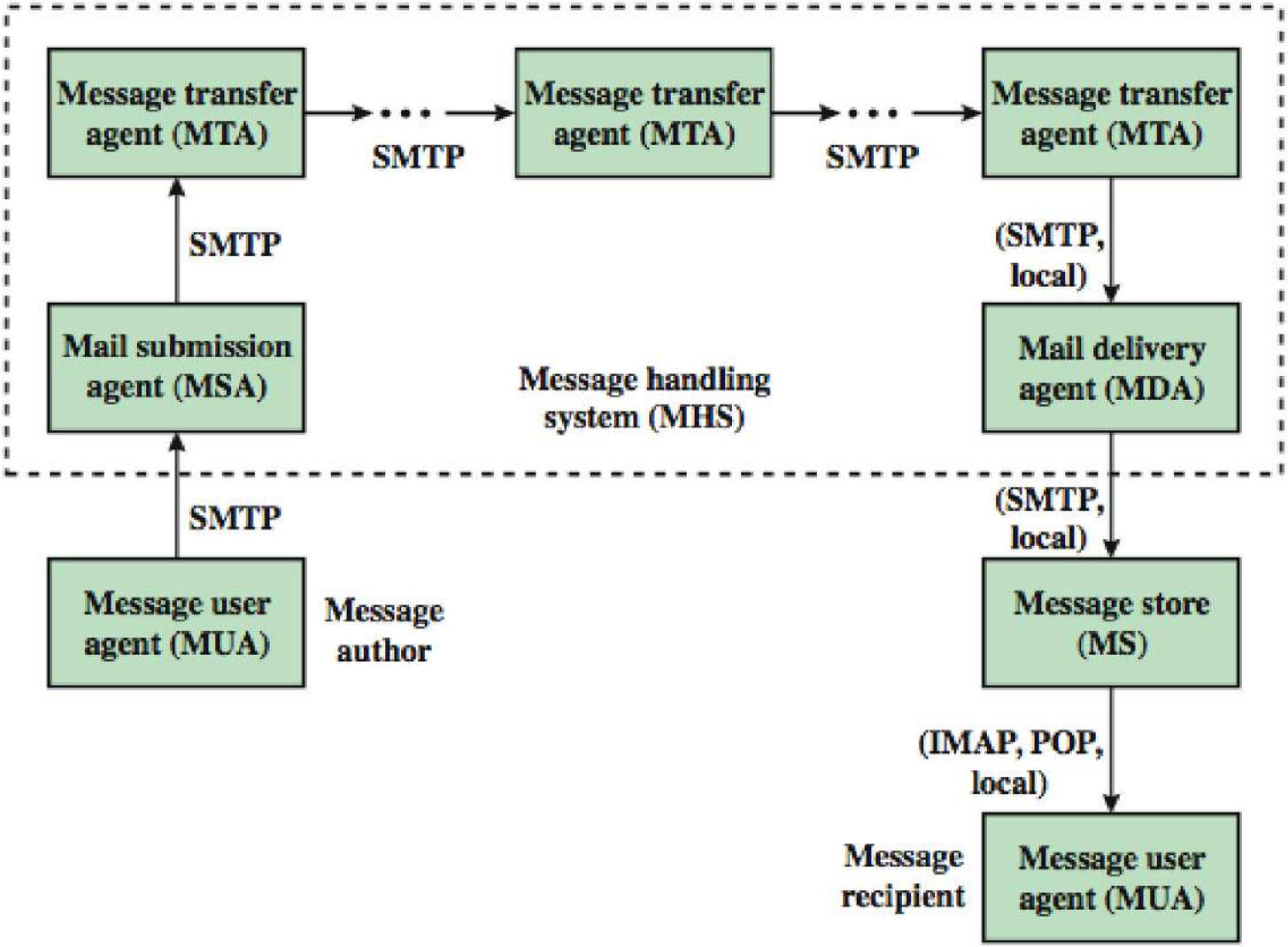
MTA

POP3/I MAP server

POP3/ IMAP

MUA
Eg: outlook gmail

Internet Mail Architecture



Internet Mail Architecture

- At fundamental the Internet mail architecture consists of a user world in the form of Message User Agents (MUA), and the transfer world, in the form of the Message Handling Service (MHS), which is composed of Message Transfer Agents (MTA).
- A MUA is usually housed in the user's computer, and referred to as a client email program.
- The Message Submission Agent (MSA) accepts the message submitted by an MUA and enforces the policies of the hosting domain and the requirements of Internet standards.
- Mail relaying is performed by a sequence of MTAs, until the message reaches a destination MDA.
- The MDA is responsible for transferring the message from the MHS to the MS.
- Typically, an MUA retrieves messages from a remote server using POP (Post Office Protocol) or IMAP (Internet Message Access Protocol)

Email Protocols

- SMTP
- STARTTLS
- Mail Access Protocols(POP3, IMAP)

Simple Mail Transfer Protocol(SMTP)

S – Sending
M – Mail
T – To
P - People

- SMTP is a text based client-server protocol.
- Client(email sender) contacts the server (next-hop recipient) and issues a set of commands to tell the server about the message to be sent.
- The transfer of message to ultimate destination can occur in following ways:
  - over a single SMTP client server or
  - an SMTP server may be intermediate relay that assumes the role of an SMTP client after receiving a message and then forwards that message to an SMTP server along a route to ultimate destination.

Simple Mail Transfer Protocol (SMTP)

- STARTTLS is a security related extension for SMTP.
- STARTTLS enables the addition of confidentiality and authentication in the exchange between SMTP agents.
- If client initiates the connection over a TLS enabled port , the server may prompt with a message indicating that STARTTLS option is available.

Mail Access Protocols

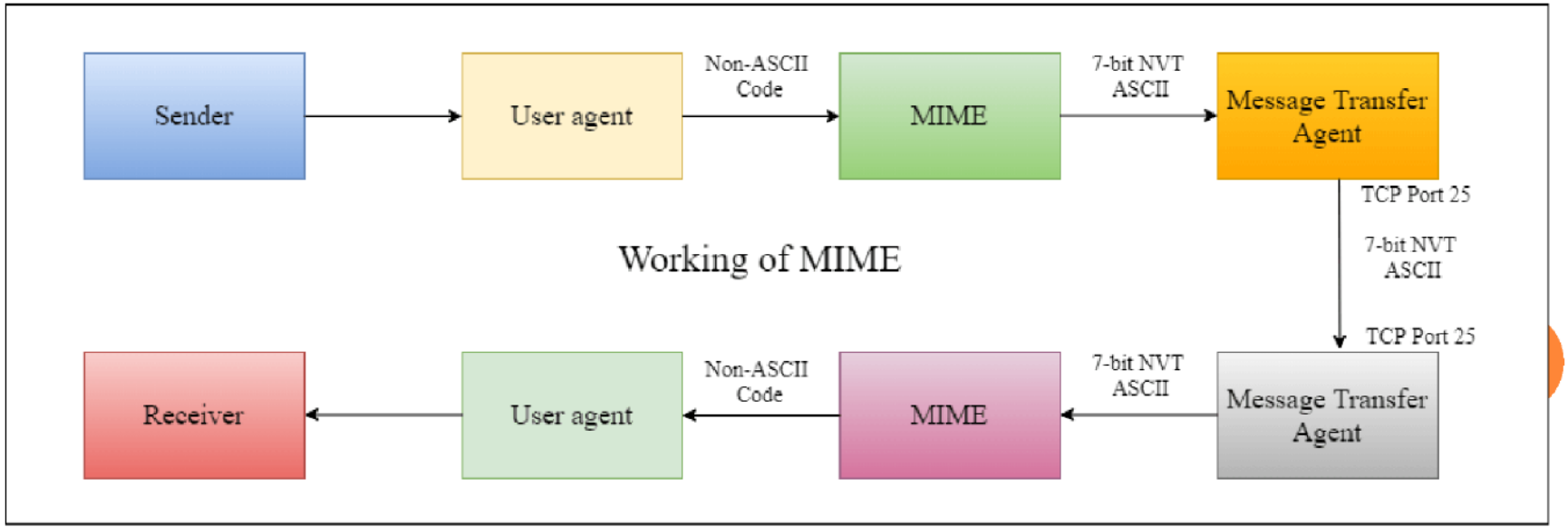
- POP3(Post office protocol): allows an email client (user agent) to download an email from email server (MTA).
- Internet Mail Access Protocol (IMAP) like POP3 enables email client to download an email from email server.
- IMAP provides more stronger authentication than POP3.

SMTP vs POP vs IMAP

SMTP	POP	IMAP
Simple Mail Transfer Protocol	Post Office Protocol	Internet Message Access Protocol
Send Mail (Push Mail)	Retrieve Mail (Pop Mail)	Retrieve Mail(Pop Mail)
Port No = 25	Port No = 110(Default) & 995 (SSL)	Port No = 143(Default) & 993(SSL)
TCP or UDP	TCP or UDP	TCP or UDP
	When it pops from Mail box and sends to receiver, IT DELETES the MAIL from MAIL BOX server.	It retains the original/ main mail in mail box only, it sends only the copy of the mail to receiver.
	Can't use in other device as its already deleted from mail box	Can be accessed at anytime and any device.

S/MIME (Secure/Multipurpose Internet Mail Extensions)

- Security enhancement to MIME email
  - original Internet RFC822 and subsequently RFC5322 and SMTP email was text only
  - MIME provided support for varying content types and multi-part messages
  - S/MIME added security enhancements
- S/MIME is supported in many mail agents
  - eg MS Outlook, Mozilla, Mac Mail etc



- Limitations of the SMTP/5322 scheme.
  - SMTP cannot transmit executable files or other binary objects.
  - SMTP cannot transmit text data that includes national language characters.
  - SMTP servers may reject mail message over a certain size and others

S/MIME Functionality

- S/MIME provides the following functions
  - Enveloped Data
    - Consists of encrypted content of any type
  - Signed Data
    - Contains a digital signature
    - The content plus signature are encoded
  - Clear-signed data
    - Contains a digital signature
    - Only signature is encoded
  - Signed and enveloped data
    - nesting of signed & encrypted entities

Functions of S/MIME

- Authentication
- Message Integrity
- Non Repudiation (using digital signatures)
- Privacy
- Data Security (using encryption)

S/MIME – User Agent Role

- Key generation
  - Generating key with RSA
- Registration
  - Register a user's public key
  - must be registered with a certification authority

S/MIME Cryptographic Algorithms

- Digital signatures: DSS & RSA
- Hash functions: SHA-1 & MD5
- Session key encryption: ElGamal & RSA
- Message encryption: AES, Triple-DES, RC2/40 and others
- MAC: HMAC with SHA-1
- have process to decide which algs to use

S/MIME Certificate Processing

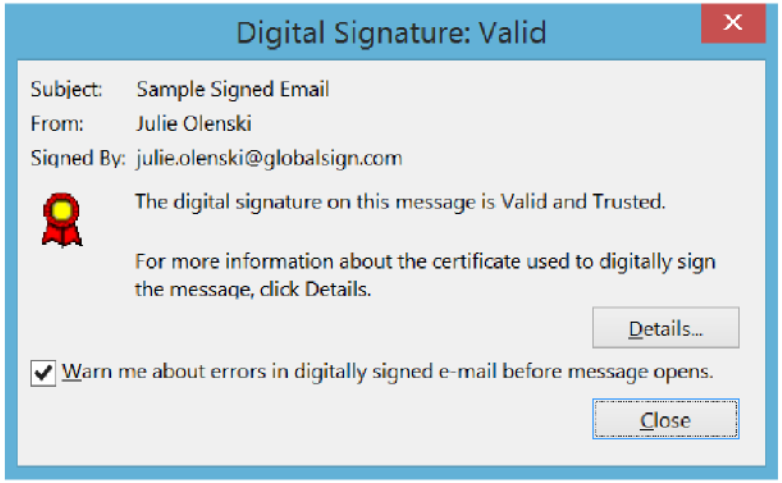
- Each client has a list of trusted CA's certs
- and own public/private key pairs & certs
- certificates must be signed by trusted CA's

Certificate Authorities

- Have several well-known CA's
- Verisign one of most widely used
- Verisign issues several types of Digital IDs
- Increasing levels of checks & hence trust
- Class 1 and Class 2 requests are processed on line, and in most cases take only a few seconds to approve. For Class 3 Digital IDs, VeriSign requires a higher level of identity assurance. An individual must prove his or her identity by providing notarized credentials or applying in person.

Digital Signature

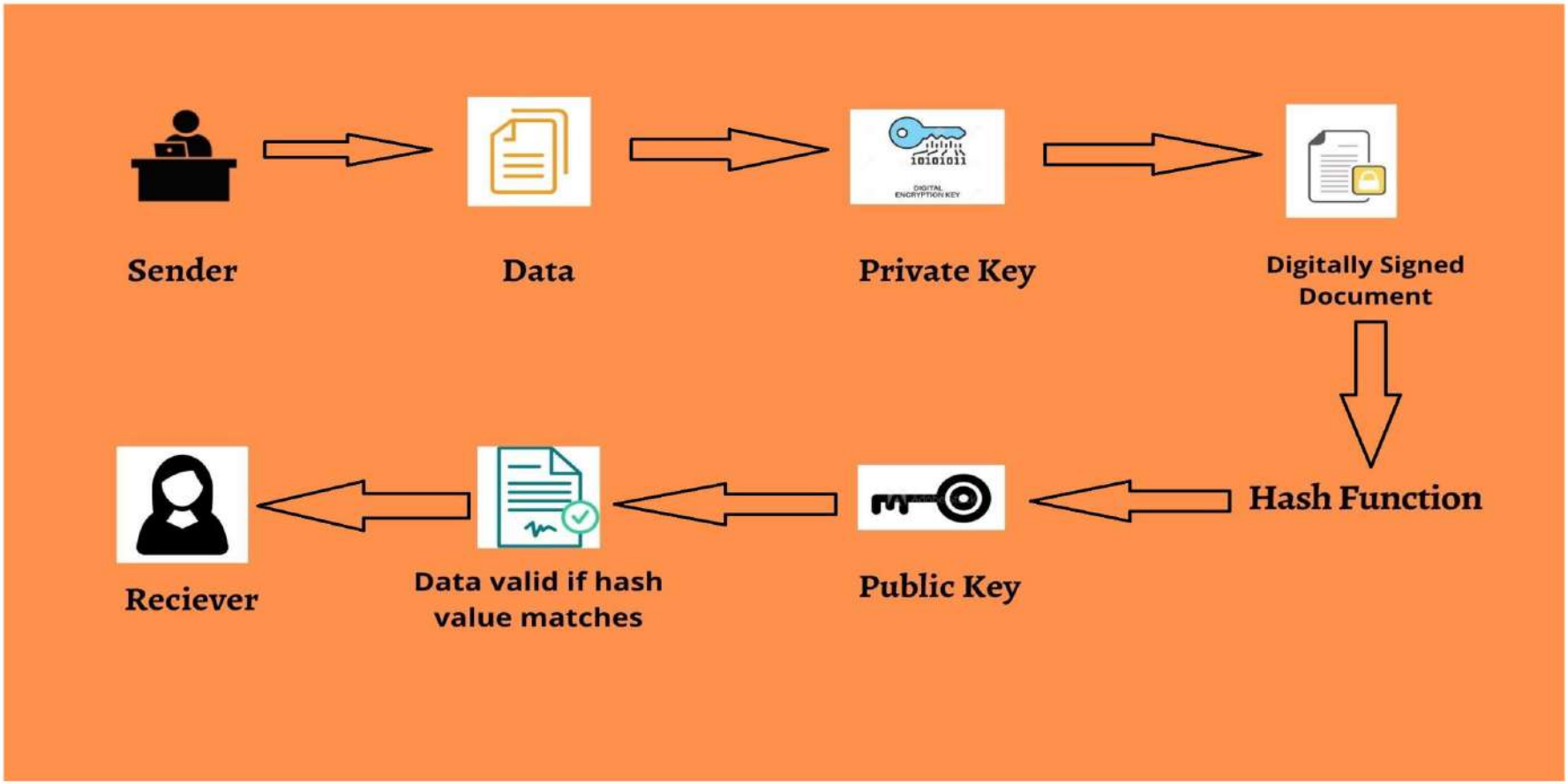
- Asymmetric Key Cryptography
- Encryption – Private Key
- Decryption – Public Key



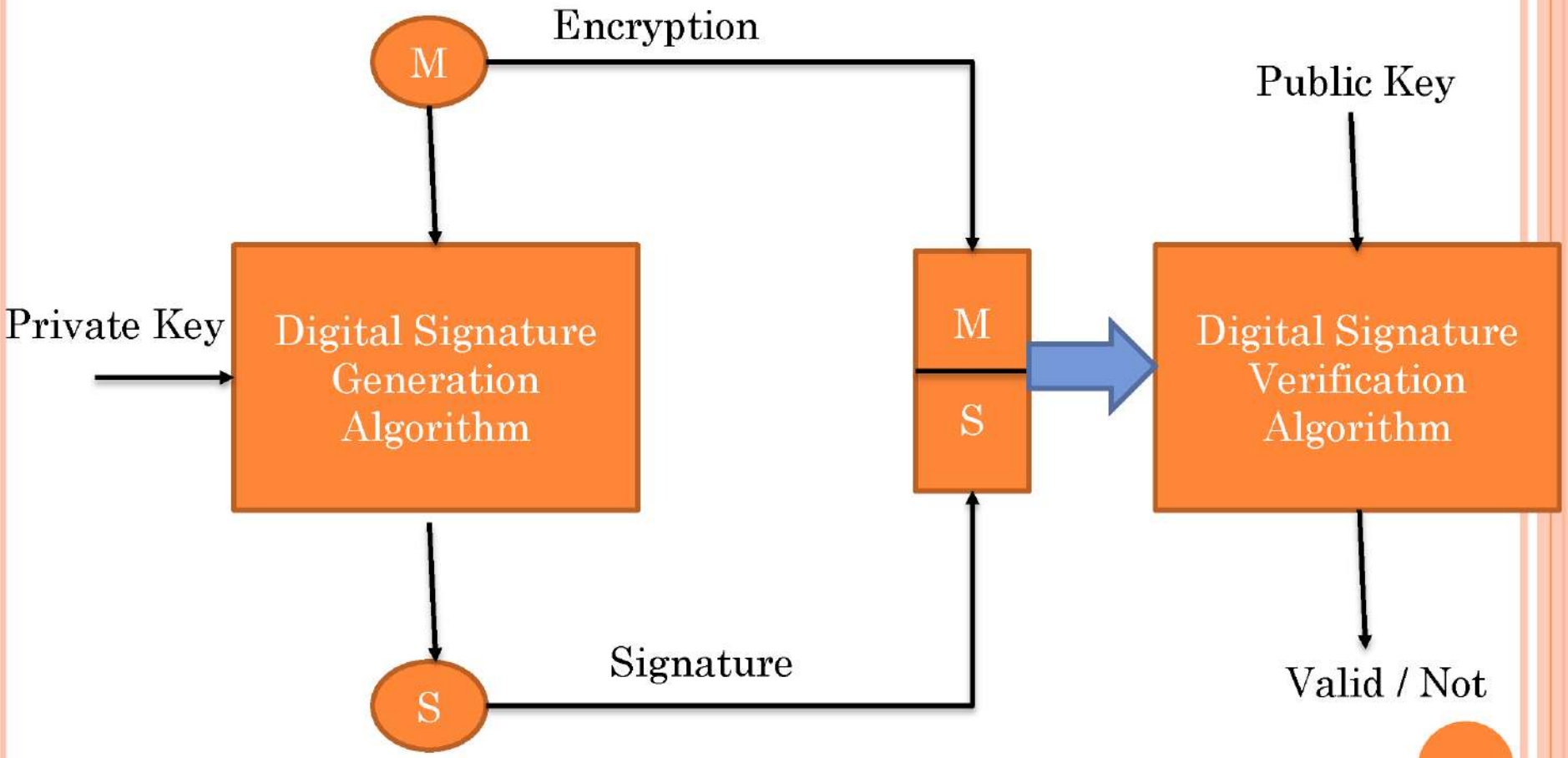
- Used for Authentication and Non-Repudiation

(Correct Person) (Cannot Deny)

- Signature: Proof of Identity (is it from correct sender /not)


Sender
Data
Private Key
Digitally Signed Document
Hash Function
Public Key
Data valid if hash value matches
Reciever

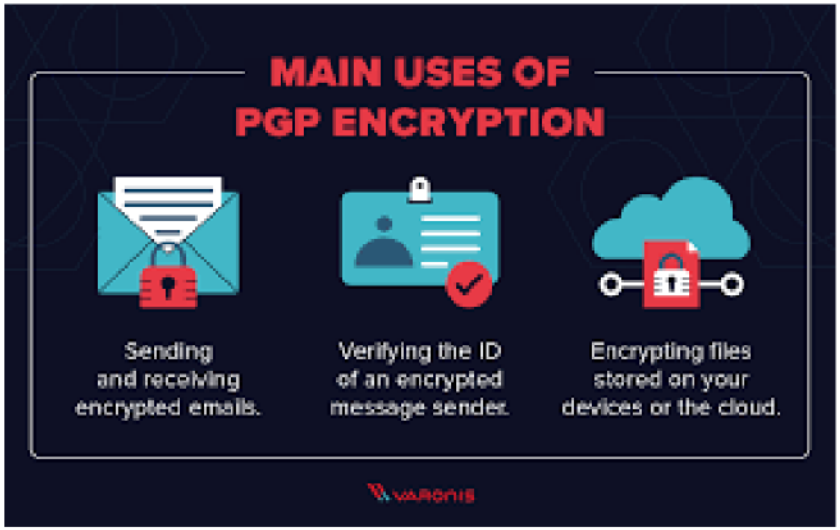
WORKING



If message matches – Valid
If message not matches – Not- Valid

Pretty Good Privacy (PGP)

- Widely used de facto secure email
- Developed by Phil Zimmermann
- Selected best available crypto algs to use
- Integrated into a single program on Unix, PC, Macintosh and other systems
- Originally free, now also have commercial versions available

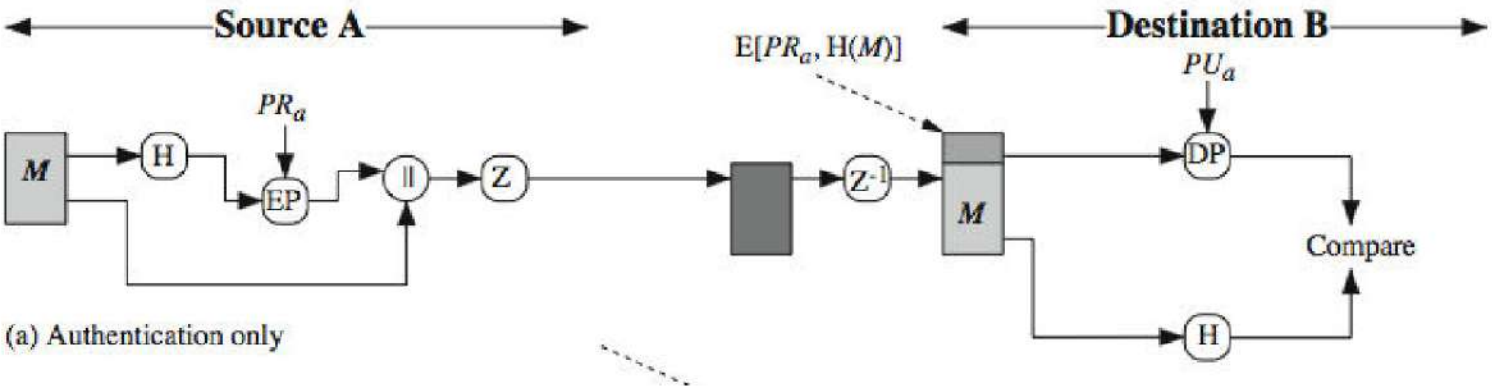


Techniques used in PGP

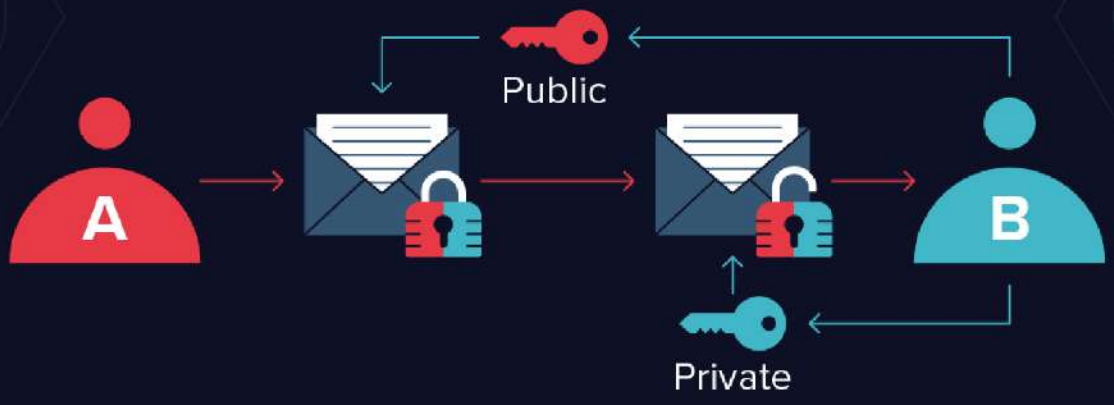
- Hashing – MD5, SHA
- Data Compression – Z
- Symmetric Key Cryptography – 1 key
- Asymmetric Key Cryptography – 2 key

PGP Operation – Authentication

1. Sender creates message
2. Make SHA-1 160-bit hash of message
3. Attach RSA signed hash to message
4. Receiver decrypts & recovers hash code
5. Receiver verifies received message hash



HOW PGP ENCRYPTION WORKS

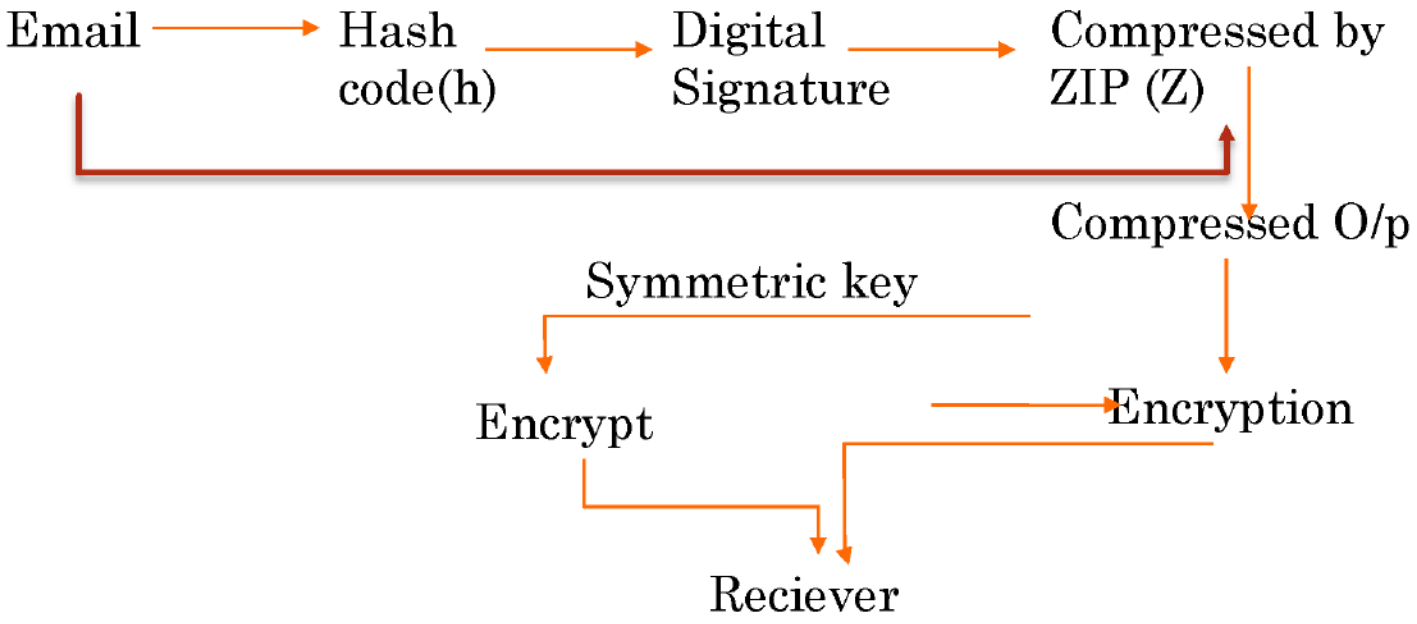


1. User A wants to send User B a private email
2. User B generates a public and private key
3. User B keeps the private key and sends back the public key
4. User A encrypts their message using the public key
5. User A sends the private encrypted message
6. User B decrypts the message with the private key

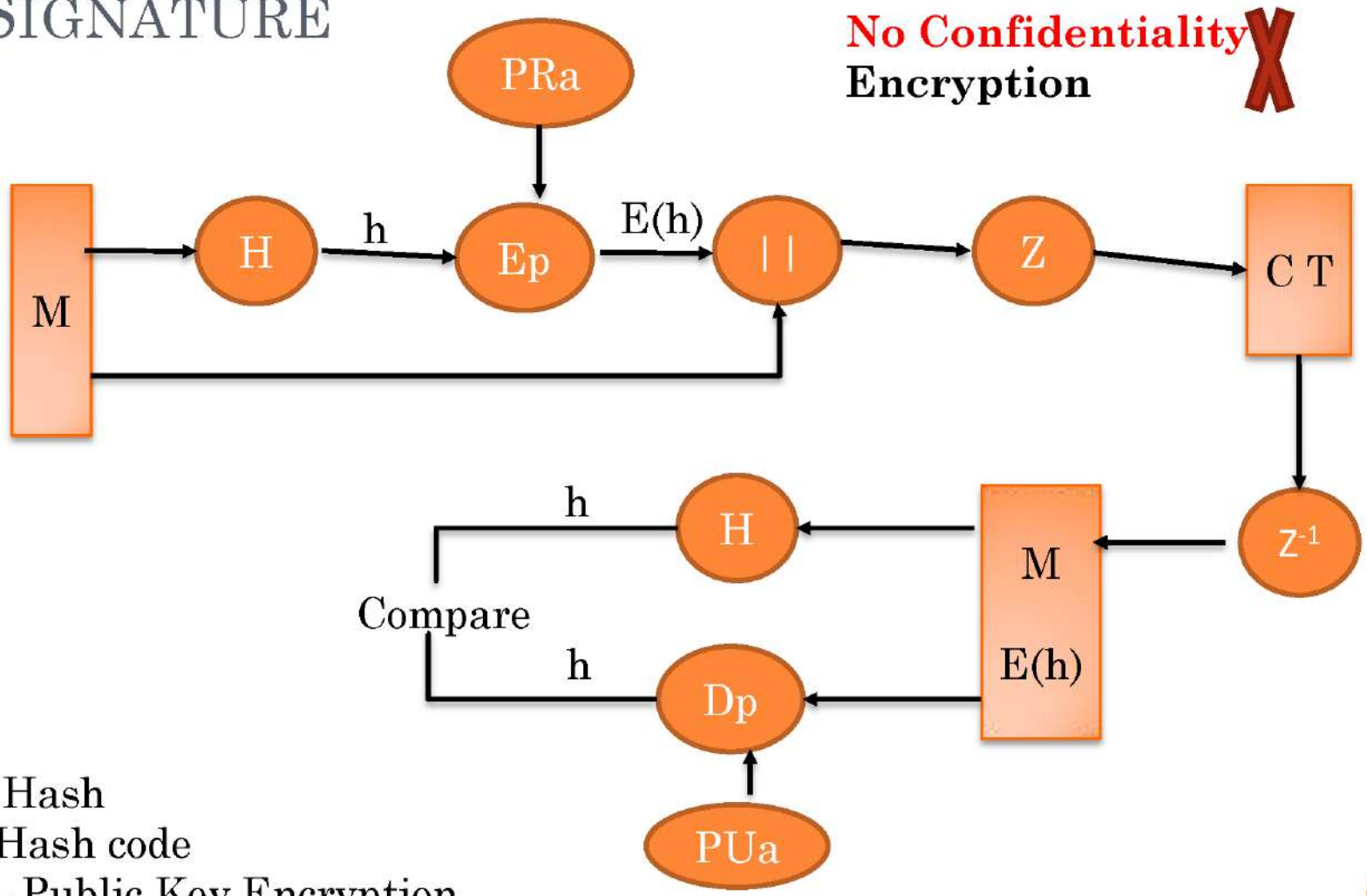
VARONIS

Services of PGP

- Case 1: Authentication
- Case 2: Confidentiality
- Case 3: Authentication + Confidentiality

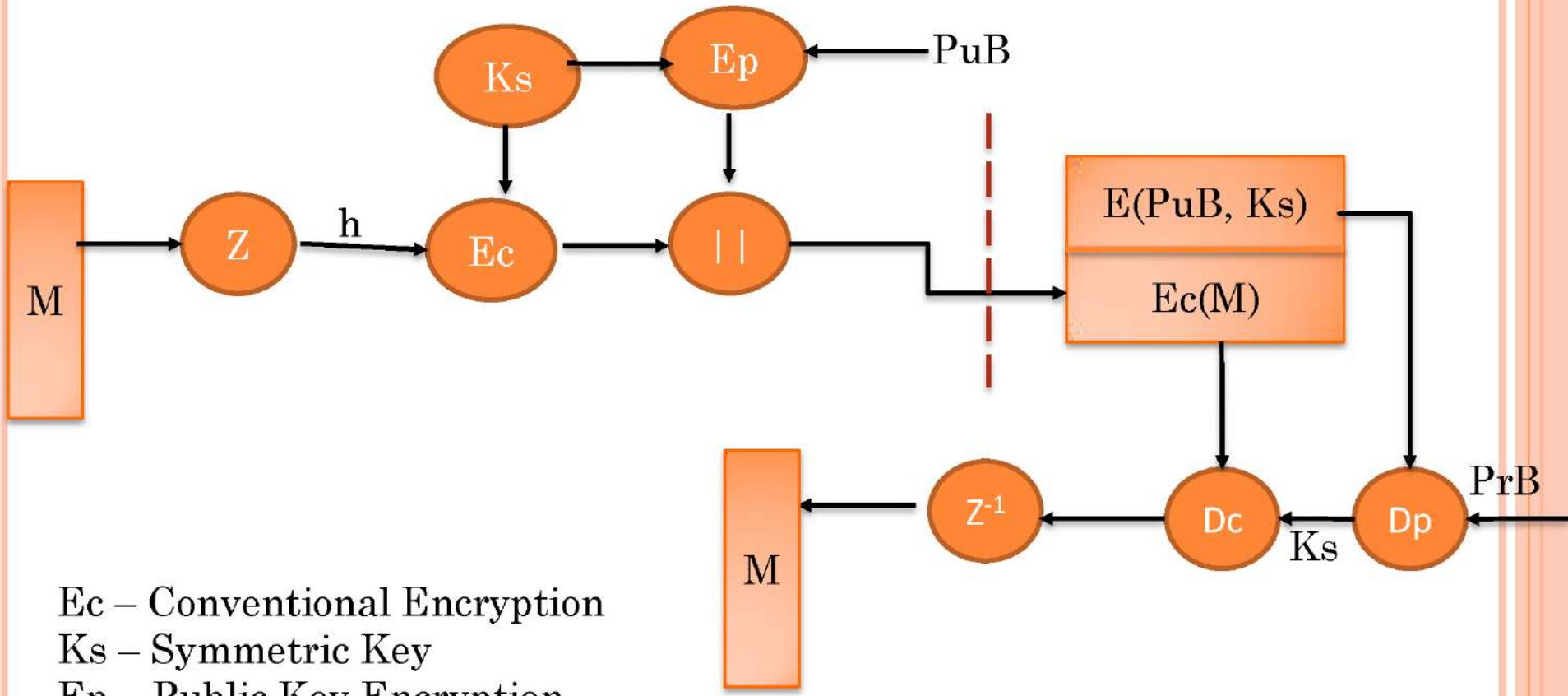


Case 1: Authentication + Digital Signature



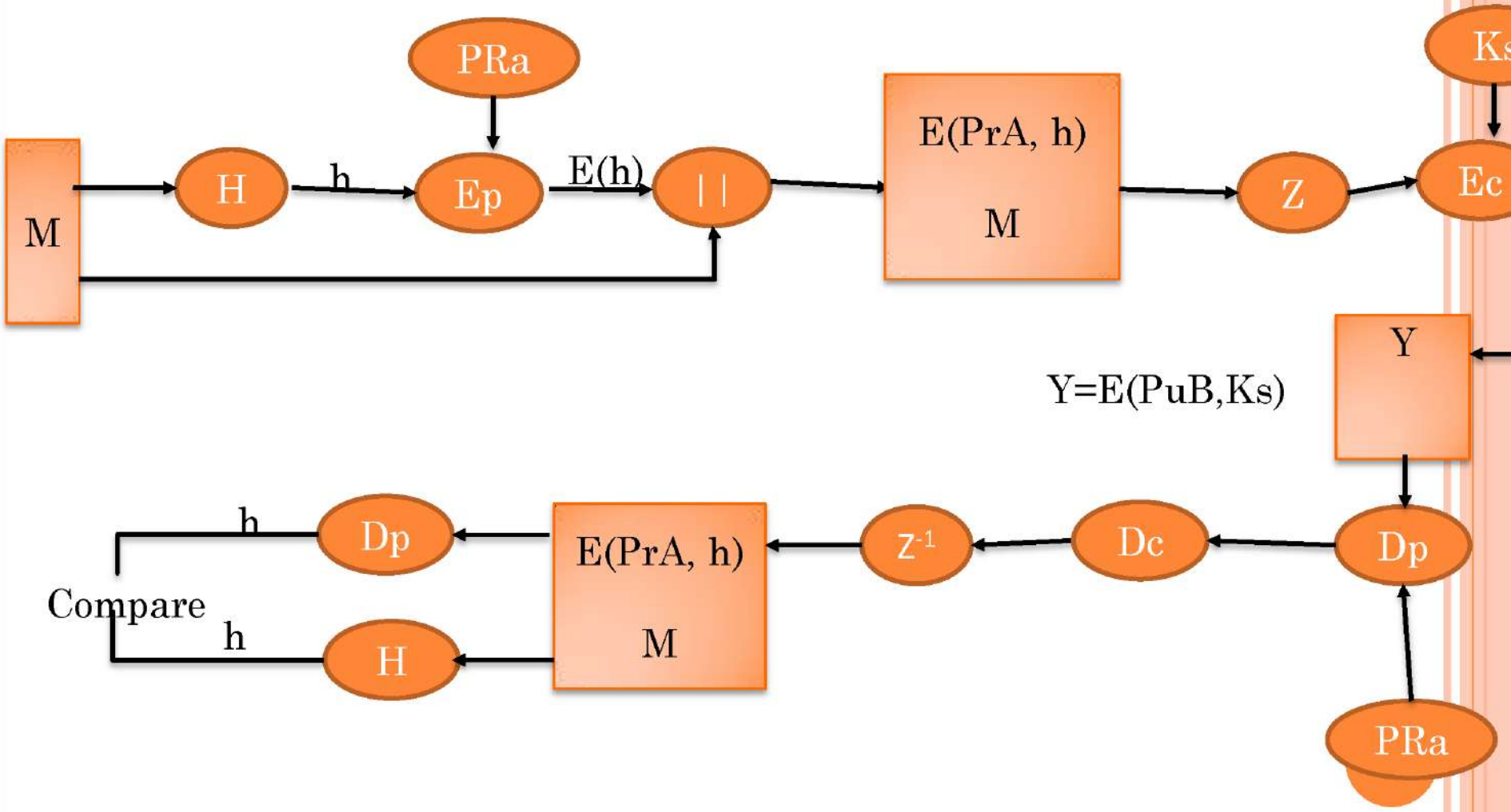
H – Hash
h – Hash code
Ep – Public Key Encryption
PrA – Private Key of A
E(h) – Encrypt of hash code

Case 2: Confidentiality



Ec – Conventional Encryption
Ks – Symmetric Key
Ep – Public Key Encryption
Dc – Conventional Decryption

Case 3: Confidentiality + Authentication



PGP Operation – Confidentiality

1. Sender forms 128-bit random session key
2. Encrypts message with session key
3. Attaches session key encrypted with RSA
4. Receiver decrypts & recovers session key
5. Session key is used to decrypt message

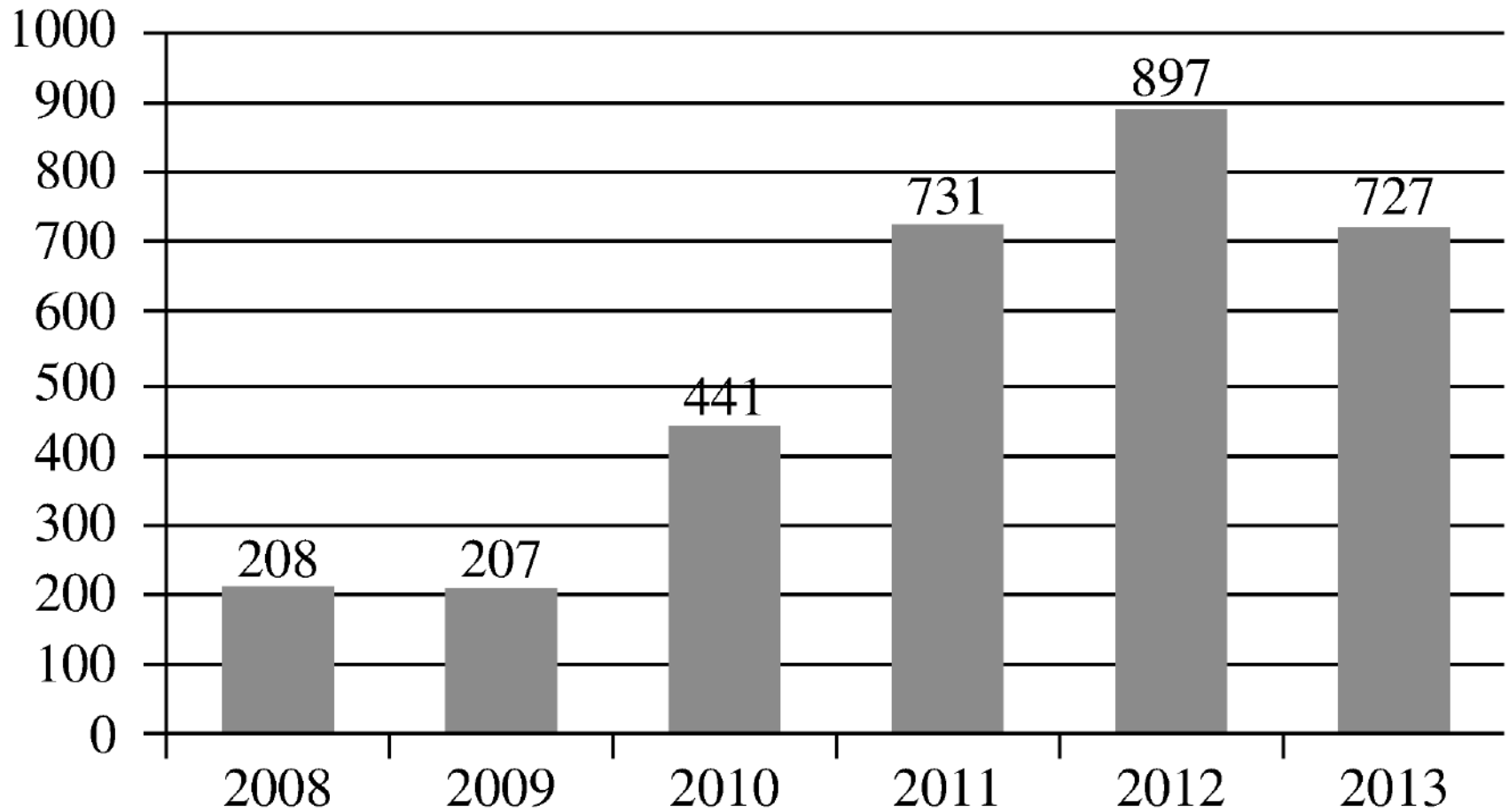
PGP Operation – Confidentiality & Authentication

- can use both services on same message
  - create signature & attach to message
  - encrypt both message & signature
  - attach RSA/ElGamal encrypted session key

PGP Operation – Compression

- by default PGP compresses message after signing but before encrypting
  - so can store uncompressed message & signature for later verification
- uses ZIP compression algorithm

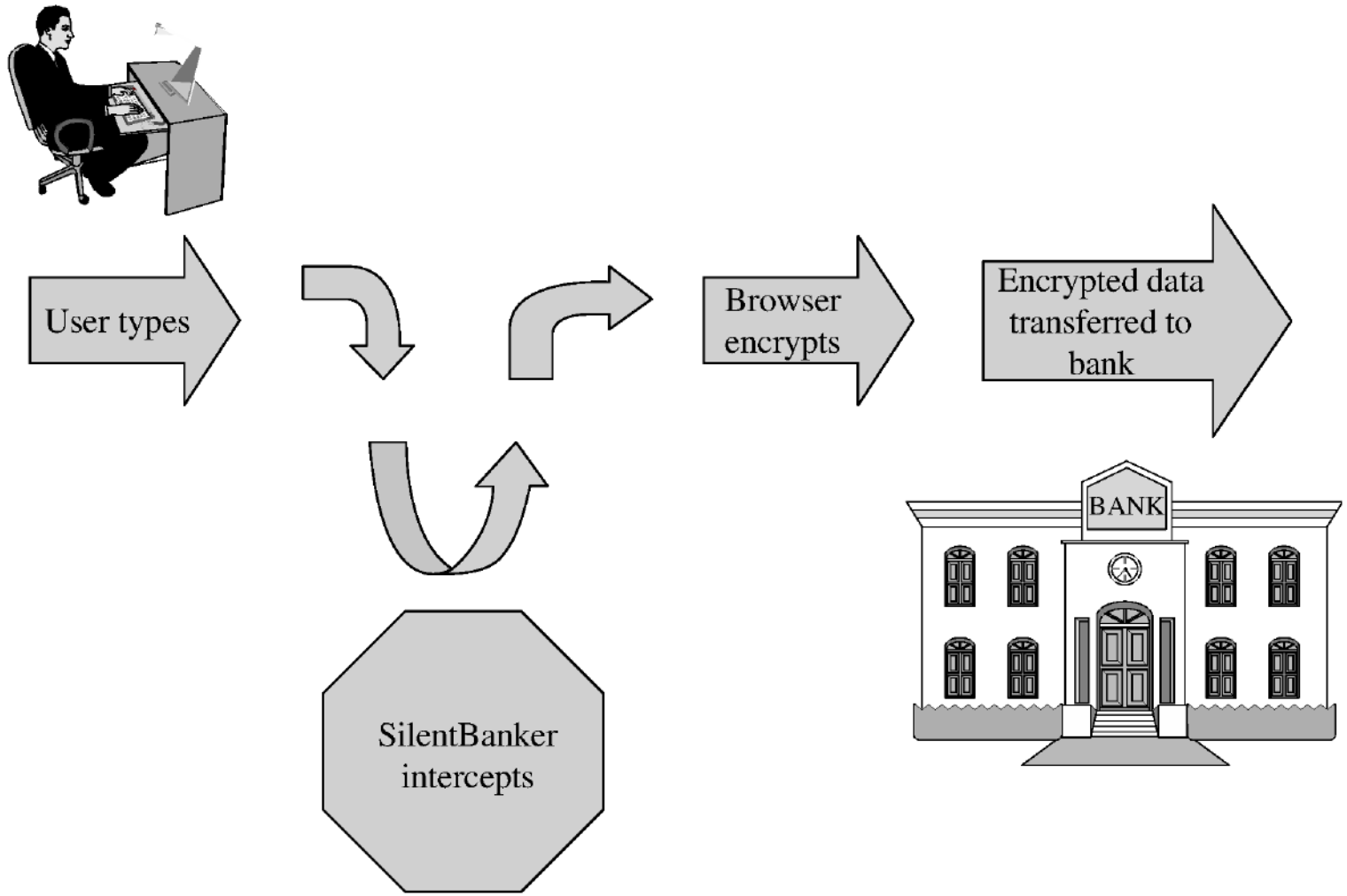
Browser Vulnerabilities



Browser Attack Types

- Man-in-the-browser
- Keystroke logger
- Page-in-the-middle
- Program download substitution
- User-in-the-middle

Man-in-the-Browser



Keystroke Logger

- Hardware or software that records all keystrokes
- May be a small dongle plugged into a USB port or can masquerade as a keyboard
- May also be installed as malware
- Not limited to browsers

Page-in-the-Middle

- User is directed to a different page than believed or intended
- Similar effect to a man-in-the-browser, where attacker can intercept and modify user input

Program Download Substitution

- Attacker creates a page with seemingly innocuous and desirable programs for download
- Instead of, or in addition to, the intended functionality, the user installs malware
- This is a very common technique for spyware

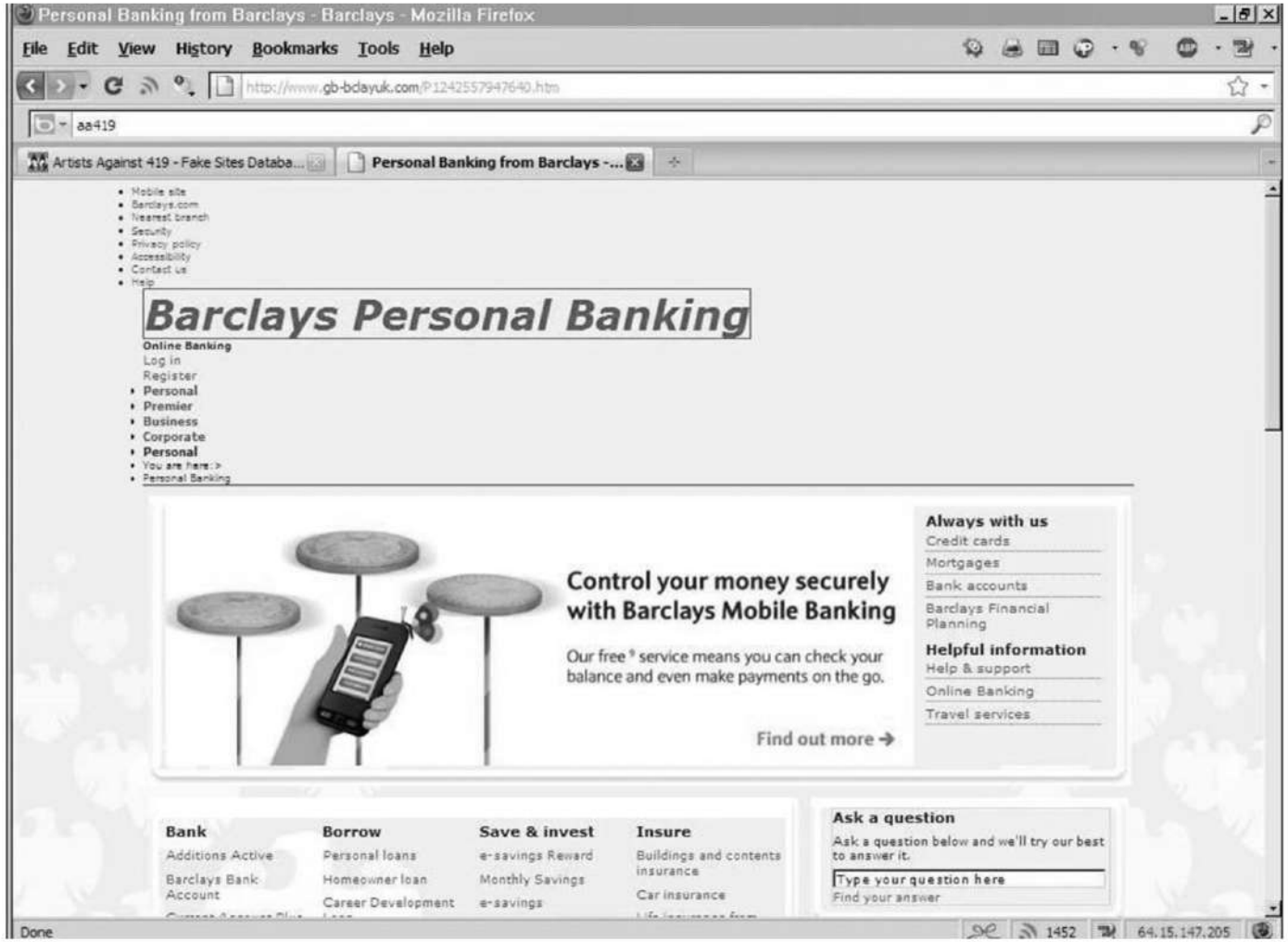
User-in-the-Middle

- Using click-bait to trick users into solving CAPTCHAs on spammers' behalf

Successful Authentication

- The attacks listed above are largely failures of authentication
- Can be mitigated with
  - Shared secret
  - One-time password
  - Out-of-band communication

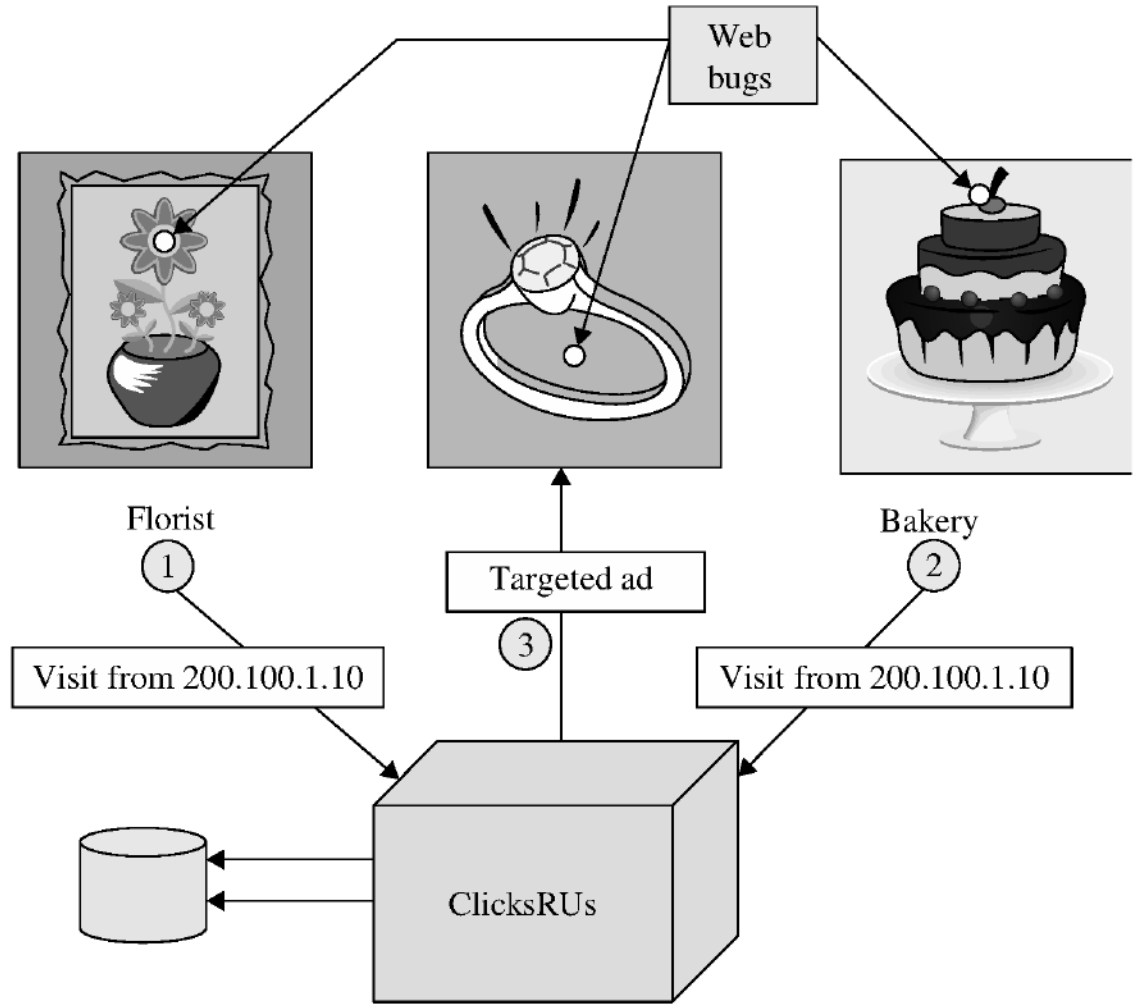
Fake Website



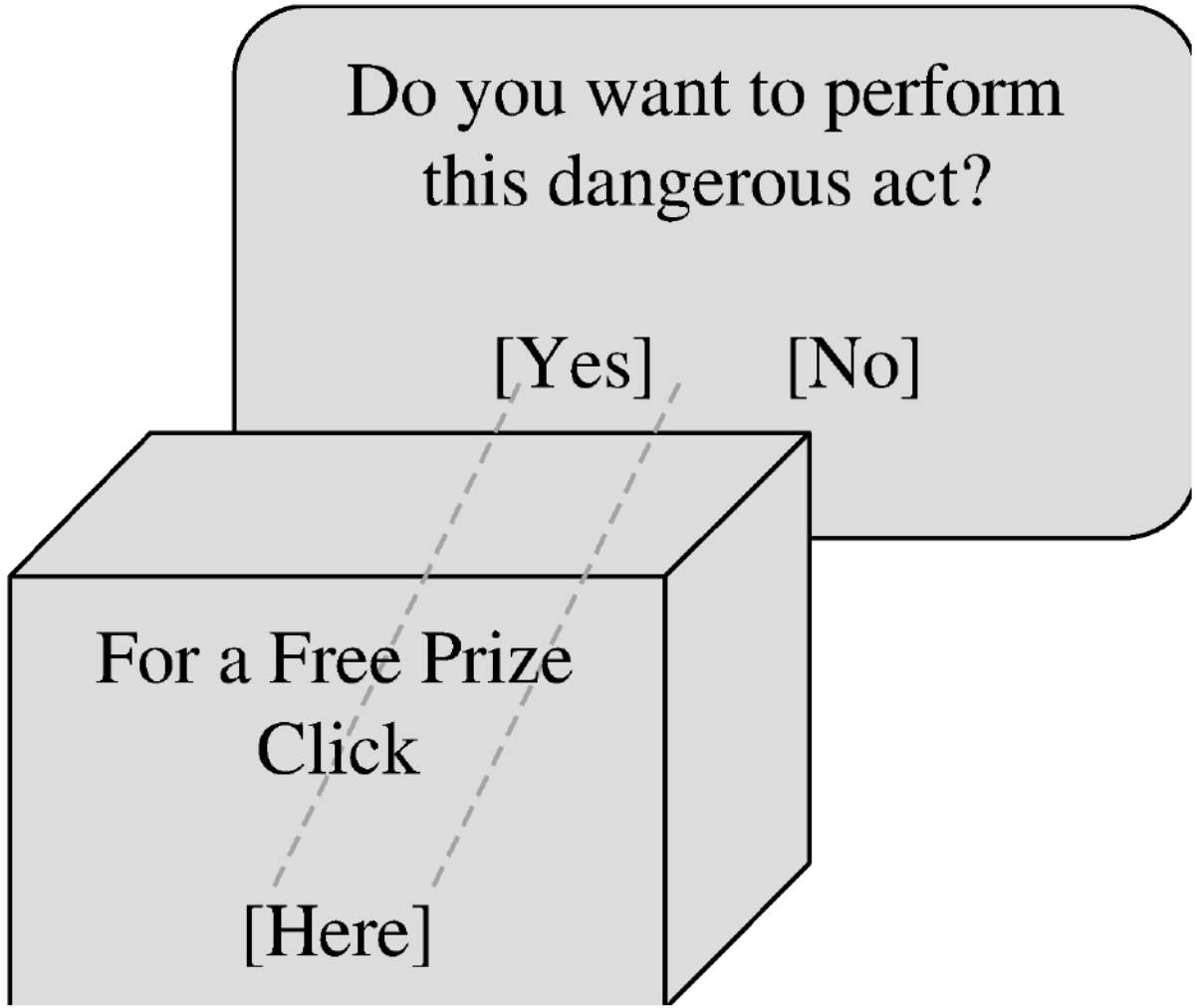
Fake Code


PDF2010
Home | Download | Members | More Info | Support
The Ultimate PDF Software Pack to
Open, Create & Edit Files
in PDF format
The BEST All in One Office Solution for your PDF files
UPDATE TO 2010 VERSION!
Top Features
50% faster than previous versions
Search & save online Internet content
Support for all Operating platforms
New and improved interface
Search single or multiple PDF files
Writer / Reader
Download the easiest software to view, create, modify and print PDF documents. The PDF format as a global exchange document format is created by Adobe and is the most efficient way to exchange information.
PDF READER WRITER PROFESSIONAL 9.0
Rated the #1 Product Online!
DOWNLOAD NOW!
FREE OFFICE SUITE INCLUDED!
Download today and receive a FREE copy of the Best ALL-IN-ONE Office Solution for Your PDF files! Get Instant access to the Ultimate Office Solution Package! Why wait, Join today and experience the most exciting PDF solution available today!
Average Rating:
Downloads: 267,927
File Size: 14.8 MB
Requirements:
Windows 2000, XP, and Vista
Compatible with all Popular Platforms
Download Now

Tracking Bug



Clickjacking



Drive-By Download

- Code is downloaded, installed, and executed on a computer without the user's knowledge
- May be the result of clickjacking, fake code, program download subsitution, etc.

Cross-Site Scripting (XSS)

- Tricking a client or server into executing scripted code by including the code in data inputs
- Scripts and HTML tags are encoded as plaintext just like user inputs, so they can take over web pages similarly to the way buffer overflow attacks can take over programs

```
Cool<br>story.<br>KCTVBigFan<script
src=http://badsite.com/xss.js></script>
```

SQL Injection

- Injecting SQL code into an exchange between an application and its database server
- Example:
  - Loading an SQL query into a variable, taking the value of acctNum from an arbitrary user input field:
  - `QUERY = "SELECT * FROM trans WHERE acct = '" + acctNum + " '; "`
  - The same query with malicious user input:
  - `QUERY = "SELECT * FROM trans WHERE acct = '2468' OR '1'='1'; "`

Dot-Dot-Slash

- Also known as “directory traversal,” this is when attackers use the term “../” to access files that are on the target web server but not meant to be accessed from outside
- Mos
```
http://yoursite.com/webhits.htw?CiWebHits&File=../../../../../winnt/system32/autoexec.nt
```
but may also be combined with other attacks, such as XSS

Server-Side Include (SSI)

- SSI is an interpreted server-side scripting language that can be used for basic web server directives, such as including files and executing commands `<!--#exec cmd="/usr/bin/telnet &"-->`
- As is the case with XSS, some websites are vulnerable to allowing users to execute SSI directives through text input

Countermeasures to Injections

- Filter and sanitize all user input
  - Need to account for every potentially valid encoding
- Make no assumptions about the range of possible user inputs—trust nothing, check everything
- Use access control mechanisms on backend servers, such as “stored procedures”

Email Spam

- Experts estimate that 60% to 90% of all email is spam
- Types of spam:
  - Advertising
    - Pharmaceuticals
    - Stocks
  - Malicious code
  - Links for malicious websites
- Spam countermeasures
  - Laws against spam exist but are generally ineffective
  - Email filters have become very effective for most spam
  - Internet service providers use volume limitations to make spammers' jobs more difficult

Phishing

- **Phishing** – Cybercriminal attempts to steal personal and financial information or infect computers and other devices with malware and viruses
  - Designed to trick you into clicking a link or providing personal or financial information
  - Often in the form of emails and websites
  - May appear to come from legitimate companies, organizations or known individuals
  - Take advantage of natural disasters, epidemics, health scares, political elections or timely events

Different forms such as:

  - **Mass Phishing** – Mass, large-volume attack intended to reach as many people as possible
  - **Whaling** – Type of spear phishing attack that targets "big fish," including high-profile individuals or those with a great deal of authority or access
  - **Clone Phishing** – Spoofed copy of a legitimate and previously delivered email, with original attachments or hyperlinks replaced with malicious versions, which is sent from a forged email address so it appears to come from the original sender or another legitimate source
  - **Advance-Fee Scam:** Requests the target to send money or bank account information to the cybercriminal
  - And **Spear Phishing…..**

Spear Phishing

- Spear phishing is on the rise because it works. Traditional security defences do not detect and stop it.
- From a cyber criminal's point of view, spear phishing is the perfect vehicle for a broad array of damaging exploits.
- Threat actors are increasingly targeting executives and other high-level employees, tricking them into activating malware that gives criminals access into their companies' environments.
- This might be ransomware that encrypts company data, then extorts fees from the victim to remediate the situation. Targeted executives are usually key leaders with titles such as chief financial officer, head of finance, senior vice president and director.
- Spear phishing emails tend to have enough detail to fool even experienced security professionals.
- A phishing campaign may blanket an entire database of email addresses, but spear phishing targets specific individuals within specific organizations with a specific mission.
- By mining social networks for personal information, an attacker can write emails that are extremely accurate and compelling.
- Once the target clicks on a link or opens an attachment, the attacker establishes a foothold in the network, enabling them to complete their illicit mission.
- 84% of organizations said a spear-phishing attack successfully penetrated their organization in 2015

Common Baiting Tactics

- **Notification from a help desk or system administrator**
  Asks you to take action to resolve an issue with your account (e.g., email account has reached its storage limit), which often includes clicking on a link and providing requested information.

- **Advertisement for immediate weight loss, hair growth or fitness prowess**
  Serves as a ploy to get you to click on a link that will infect your computer or mobile device with malware or viruses.

- **Attachment labeled "invoice" or "shipping order"**
  Contains malware that can infect your computer or mobile device if opened. May contain what is known as "ransomware," a type of malware that will delete all files unless you pay a specified sum of money.

- **Notification from what appears to be a credit card company**
  Indicates someone has made an unauthorized transaction on your account. If you click the link to log in to verify the transaction, your username and password are collected by the scammer.

- **Fake account on a social media site**
  Mimics a legitimate person, business or organization. May also appear in the form of an online game, quiz or survey designed to collect information from your account.

Spear Phishing Characteristics

A spear-phishing attack can display one or more of the following characteristics:

• Blended or multi-vector threat. Spear phishing uses a blend of email spoofing, dynamic URLs and drive-by downloads to bypass traditional defenses.

• Use of zero-day vulnerabilities. Advanced spear-phishing attacks leverage zero-day vulnerabilities in browsers, plug-ins and desktop applications to compromise systems.

• Multi-stage attack. The initial exploit of systems is the first stage of an APT attack that involves further stages of malware outbound communications, binary downloads and data exfiltration.

• Well-crafted email forgeries: Spearphishing email threats are usually targeted to individuals, so they don’t bear much resemblance to the high-volume, broadcast spam that floods the Internet. This means traditional reputation and spam filters routinely miss these messages, rendering traditional email protections ineffective.

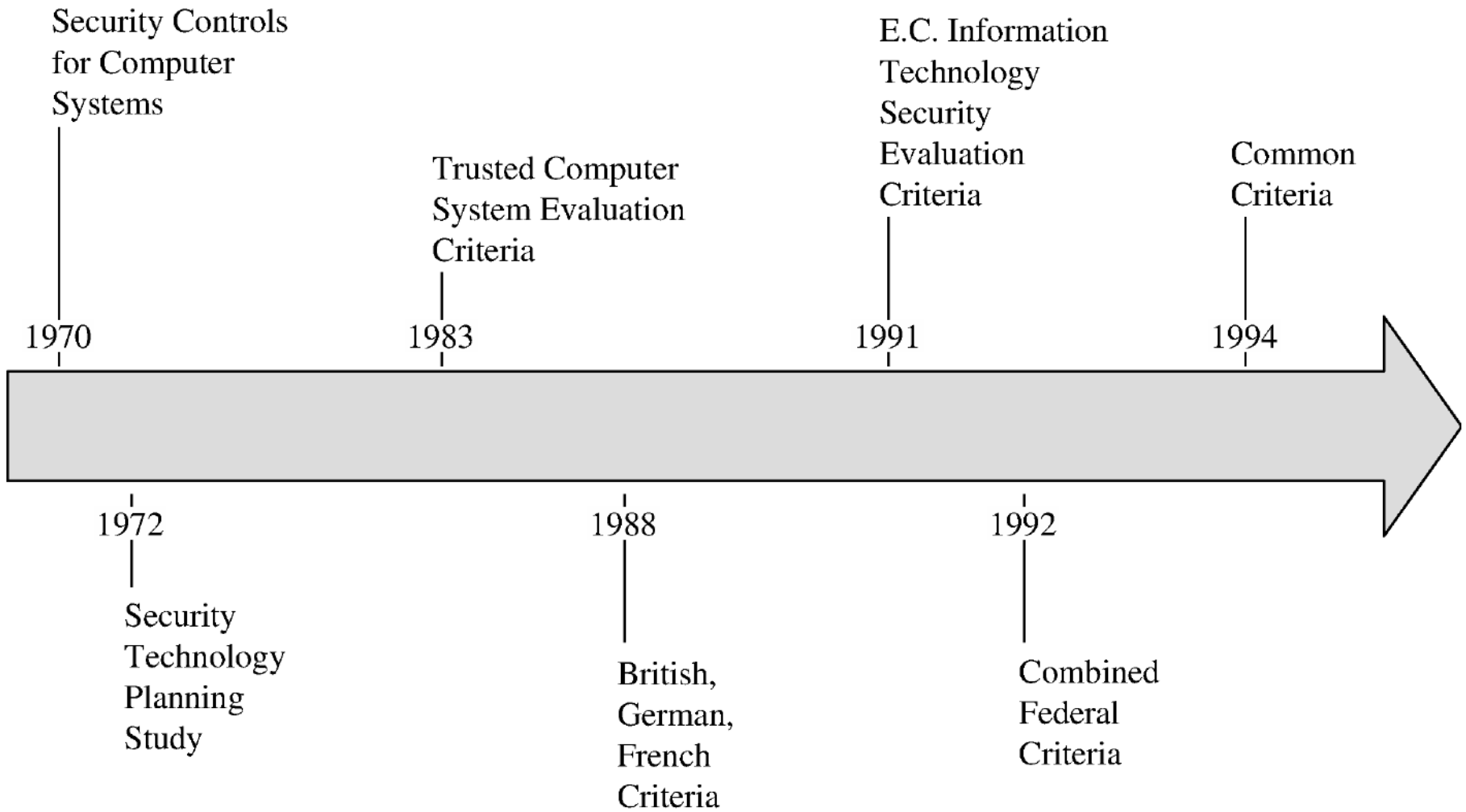
How to protect against phishing

- STOP. THINK. CONNECT.
  - Before you click, look for common baiting tactics e.g. Requests for personal information, Announcement indicating you won a prize or lottery or Requests for donations
  - Look for spelling errors (e.g., “pessward”), lack of punctuation or poor grammar
  - Hyperlinked URL differs from the one displayed, or it is hidden
  - Threatening language that calls for immediate action
- Install and maintain antivirus software on your electronic devices
- Use email filters to reduce spam and malicious traffic
- Be wary of messages asking for passwords or other personal information
  - All reputable businesses and organizations will never ask for your password via email
- Never send passwords, bank account numbers or other private information in an email
  - Do not reply to requests for this information
  - Verify by contacting the company or individual, but do not use the contact information included in the message
- Do not click on any hyperlinks in the email
  - User your computer mouse to hover over each link to verify its actual destination, even if the message appears to be from a trusted source
  - Pay attention to the URL and look for a variation in spelling or different domain (e.g., ulster.ac vs. ulster.com)
  - Consider navigating to familiar sites on your own instead of using links within messages
- Examine websites closely
  - Malicious websites may look identical to legitimate sites
  - Look for “https://” or a lock icon in the address bar before entering any sensitive information on a website

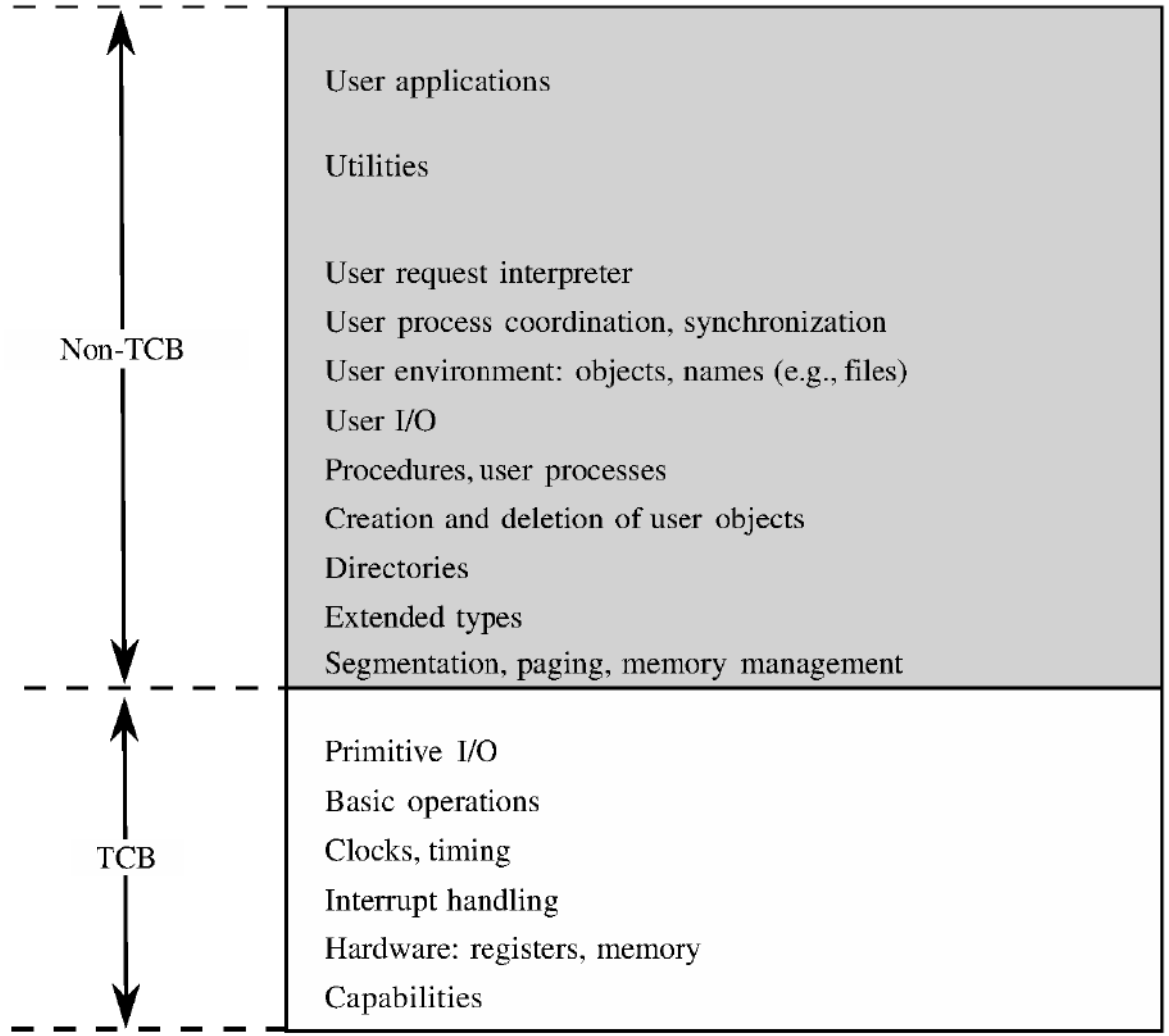
Trusted Systems

- A trusted system is one that has been shown to warrant some degree of trust that it will perform certain activities faithfully, that is, in accordance with users' expectations.
- Characteristics of a trusted system:
  - A defined policy that details what security qualities it enforces
  - Appropriate measures and mechanisms by which it can enforce security adequately
  - Independent scrutiny or evaluation to ensure that the mechanisms have been selected and implemented properly

History of Trusted Systems



Trusted Computing Base (TCB)



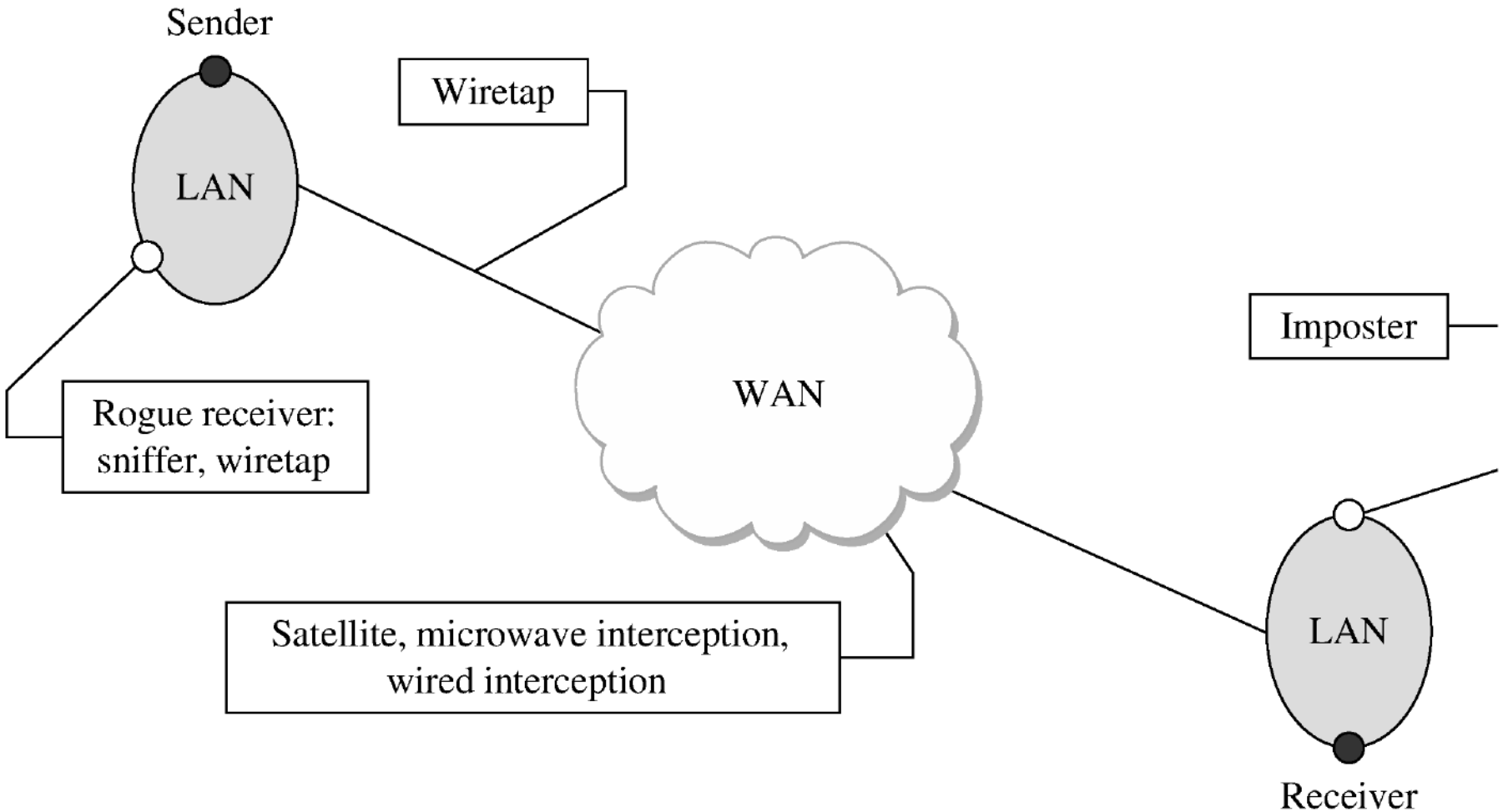
Other Trusted System Characteristics

- Secure startup
  - System startup is a tricky time for security, as most systems load basic I/O functionality before being able to load security functions
- Trusted path
  - An unforgeable connection by which the user can be confident of communicating directly with the OS
- Object reuse control
  - OS clears memory before reassigning it to ensure that leftover data doesn't become compromised
- Audit
  - Trusted systems track security-relevant changes, such as installation of new programs or OS modification
  - Audit logs must be protected against tampering and deletion

Network Transmission Media

- Cable
- Optical fiber
- Microwave
- Wi-Fi
- Satellite communication

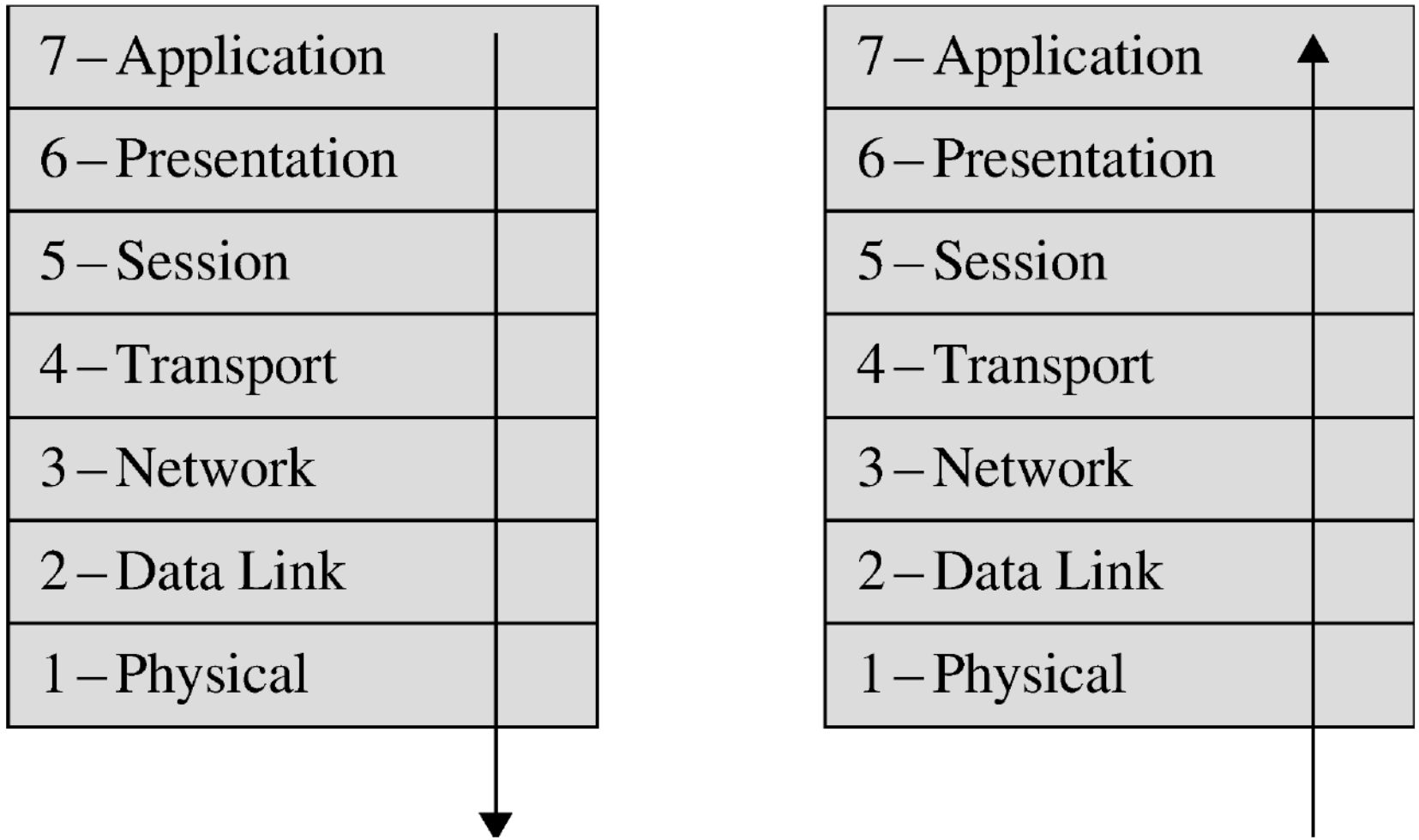
Communication Media Vulnerability



Communication Media Pros/Cons

Medium	Strengths	Weaknesses
Wire	• Widely used • Inexpensive to buy, install, maintain	• Susceptible to emanation • Susceptible to physical wiretapping
Optical fiber	• Immune to emanation • Difficult to wiretap	• Potentially exposed at connection points
Microwave	• Strong signal, not seriously affected by weather	• Exposed to interception along path of transmission • Requires line of sight location • Signal must be repeated approximately every 30 miles (50 kilometers)
Wireless (radio, WiFi)	• Widely available • Built into many computers	• Signal degrades over distance; suitable for short range • Signal interceptable in circular pattern around transmitter
Satellite	• Strong, fast signal	• Delay due to distance signal travels up and down • Signal exposed over wide area at receiving end

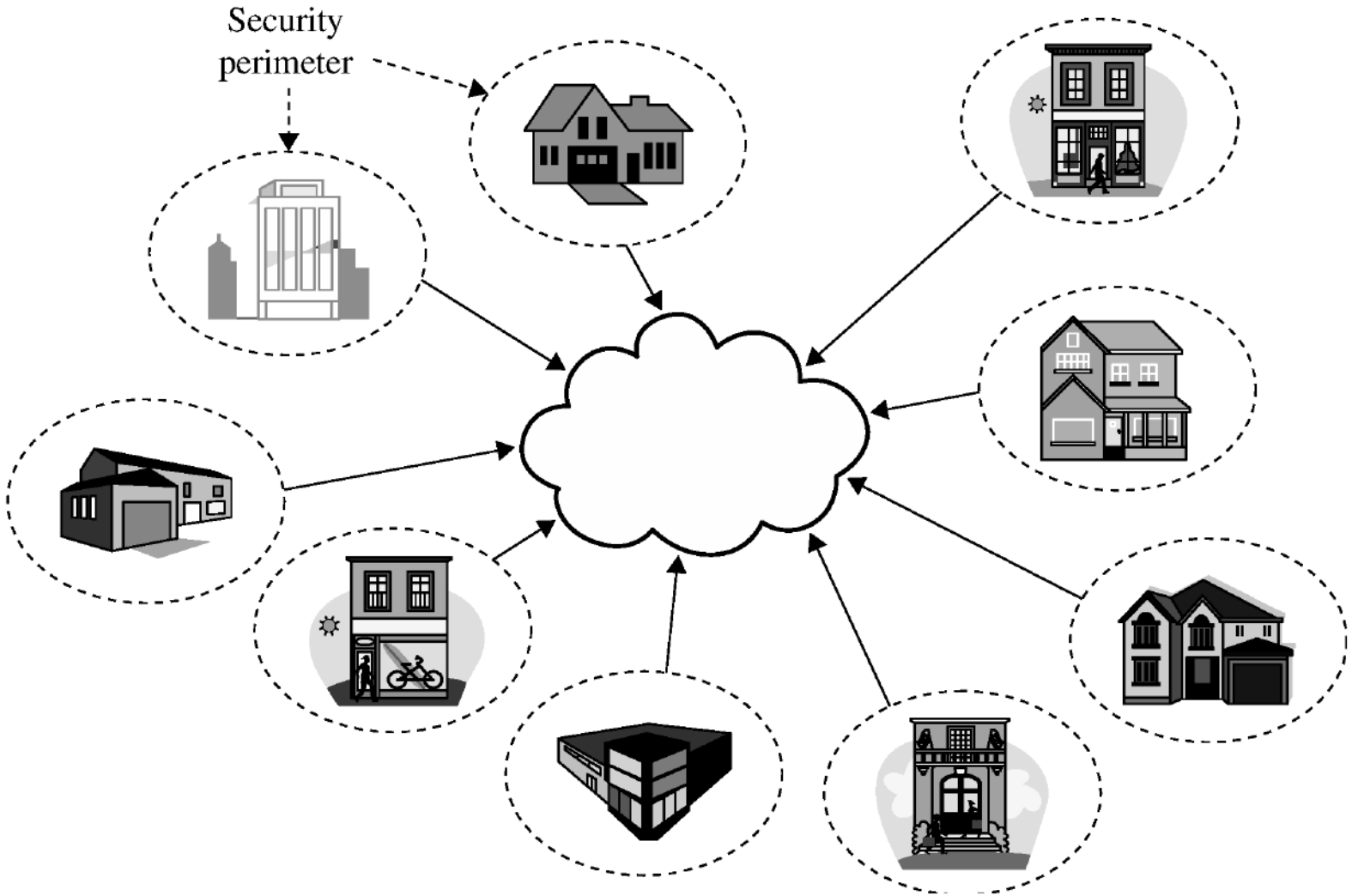
The OSI Model



Threats to Network Communications

- *Interception*, or unauthorized viewing
- *Modification*, or unauthorized change
- *Fabrication*, or unauthorized creation
- *Interruption*, or preventing authorized access

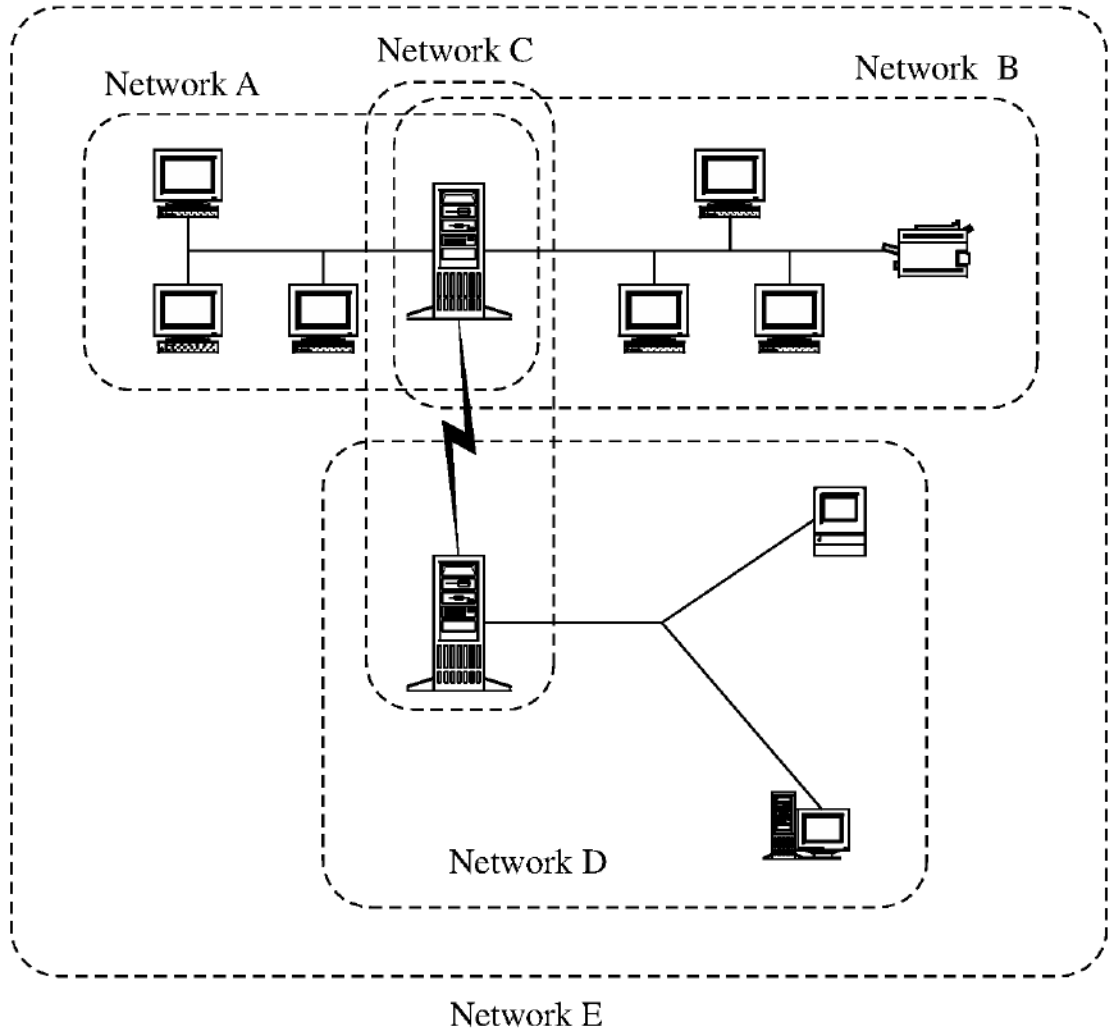
Security Perimeters



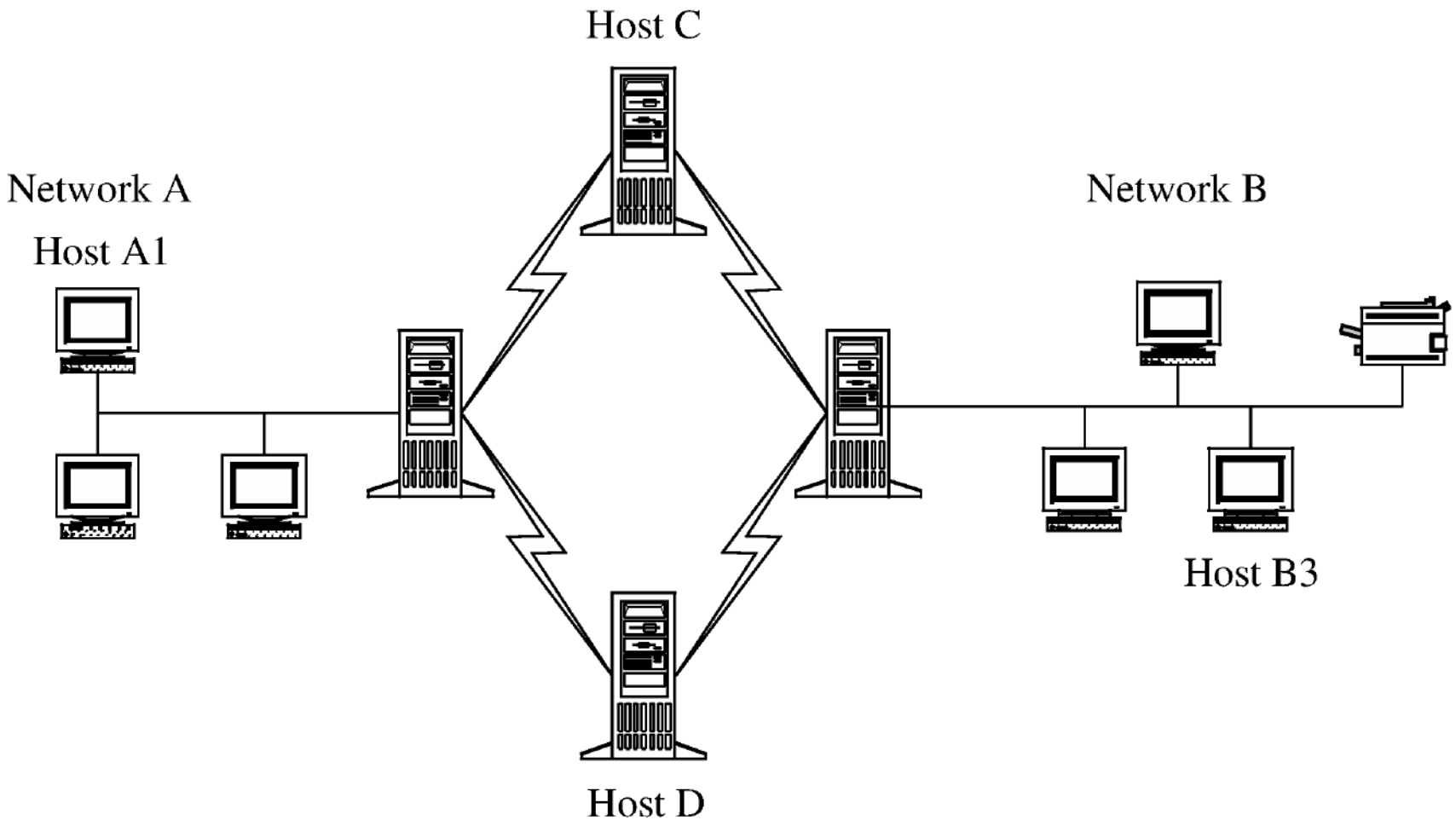
What Makes a Network Vulnerable to Interception?

- Anonymity
  - An attacker can attempt many attacks, anonymously, from thousands of miles away
- Many points of attack
  - Large networks mean many points of potential entry
- Sharing
  - Networked systems open up potential access to more users than do single computers
- System complexity
  - One system is very complex and hard to protect; networks of many different systems, with disparate OSs, vulnerabilities, and purposes are that much more complex
- Unknown perimeter
  - Networks, especially large ones, change all the time, so it can be hard to tell which systems belong and are behaving, and impossible to tell which systems bridge networks
- Unknown path
  - There may be many paths, including untrustworthy ones, from one host to another

Unknown Perimeter



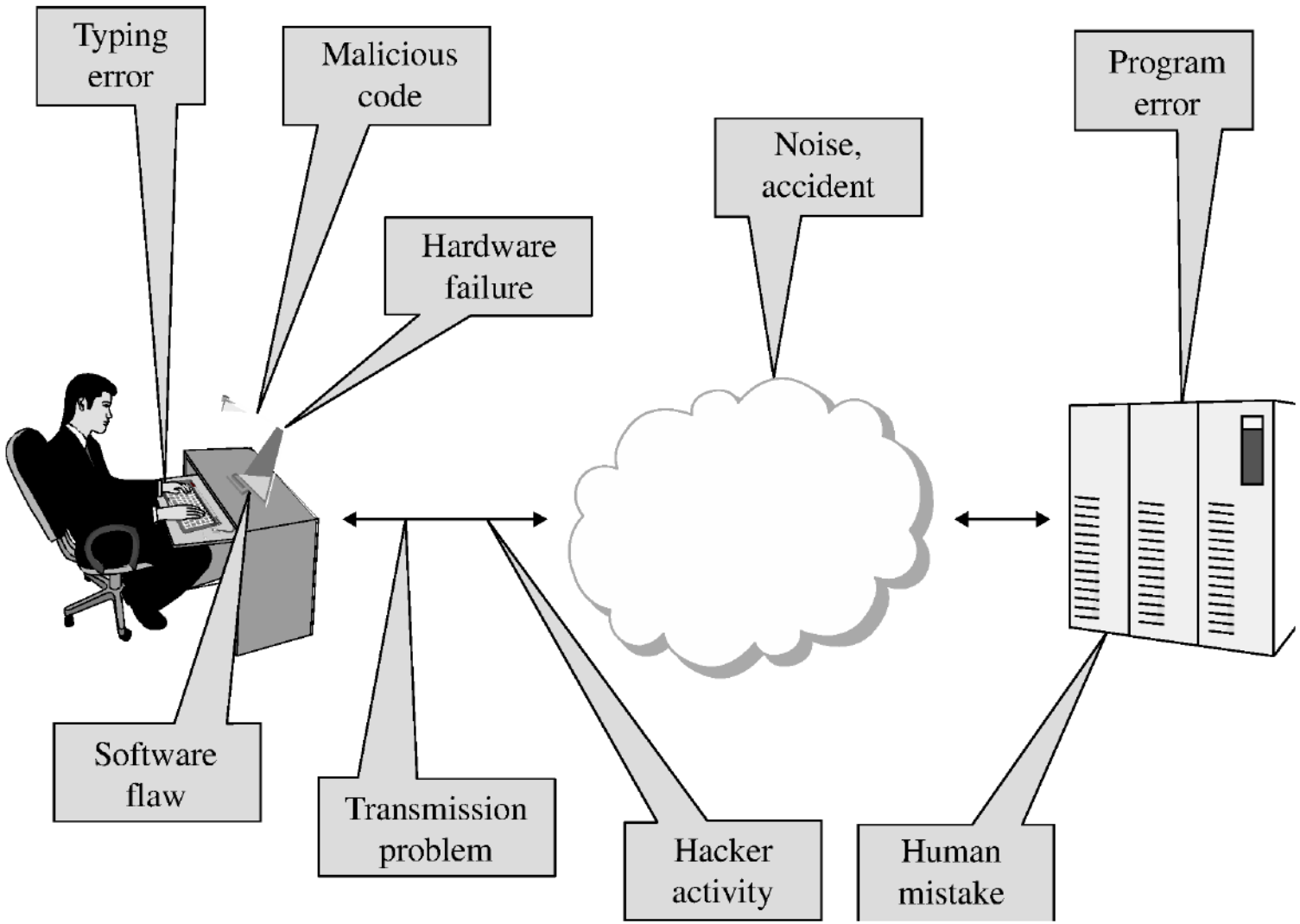
Unknown Path



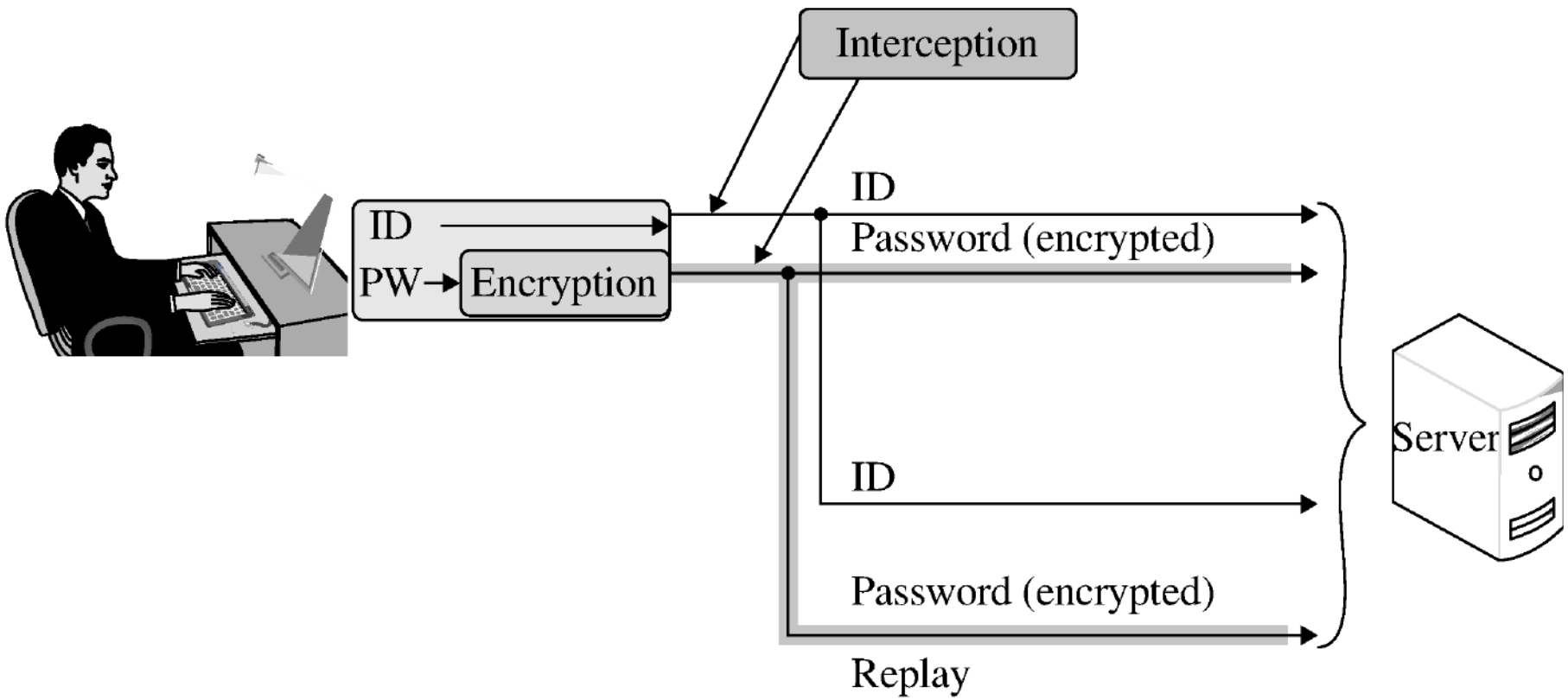
Modification and Fabrication

- Data corruption
  - May be intentional or unintentional, malicious or nonmalicious, directed or random
- Sequencing
  - Permuting the order of data, such as packets arriving in sequence
- Substitution
  - Replacement of one piece of a data stream with another
- Insertion
  - A form of substitution in which data values are inserted into a stream
- Replay
  - Legitimate data are intercepted and reused

Sources of Data Corruption



Simple Replay Attack



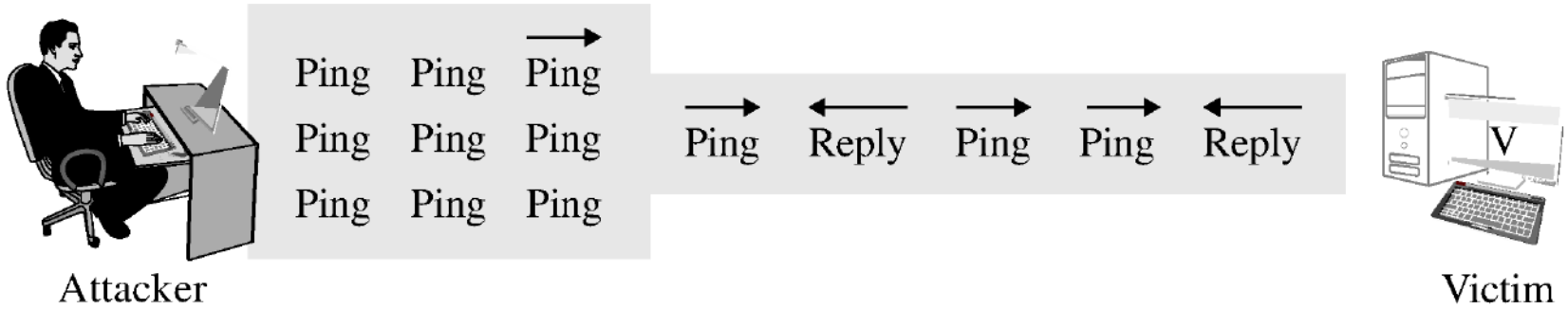
Interruption: Loss of Service

- Routing
  - Internet routing protocols are complicated, and one misconfiguration can poison the data of many routers
- Excessive demand
  - Network capacity is finite and can be exhausted; an attacker can generate enough demand to overwhelm a critical part of a network
- Component failure
  - Component failures tend to be sporadic and unpredictable, and will cause loss of service if not planned for

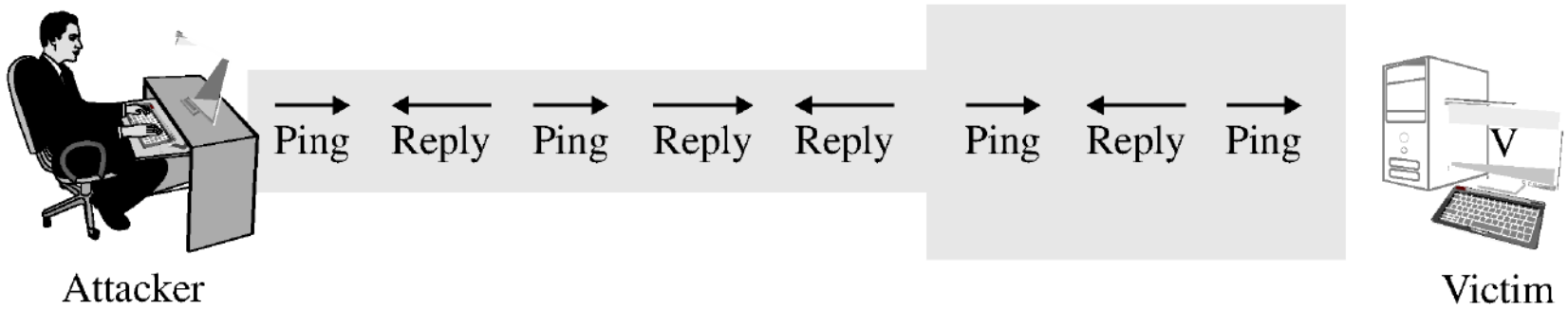
Denial of Service (DoS)

- DoS attacks are attempts to defeat a system's availability
- Volumetric attacks
- Application-based attacks
- Disabled communications
- Hardware or software failure

DoS Attack: Ping Flood

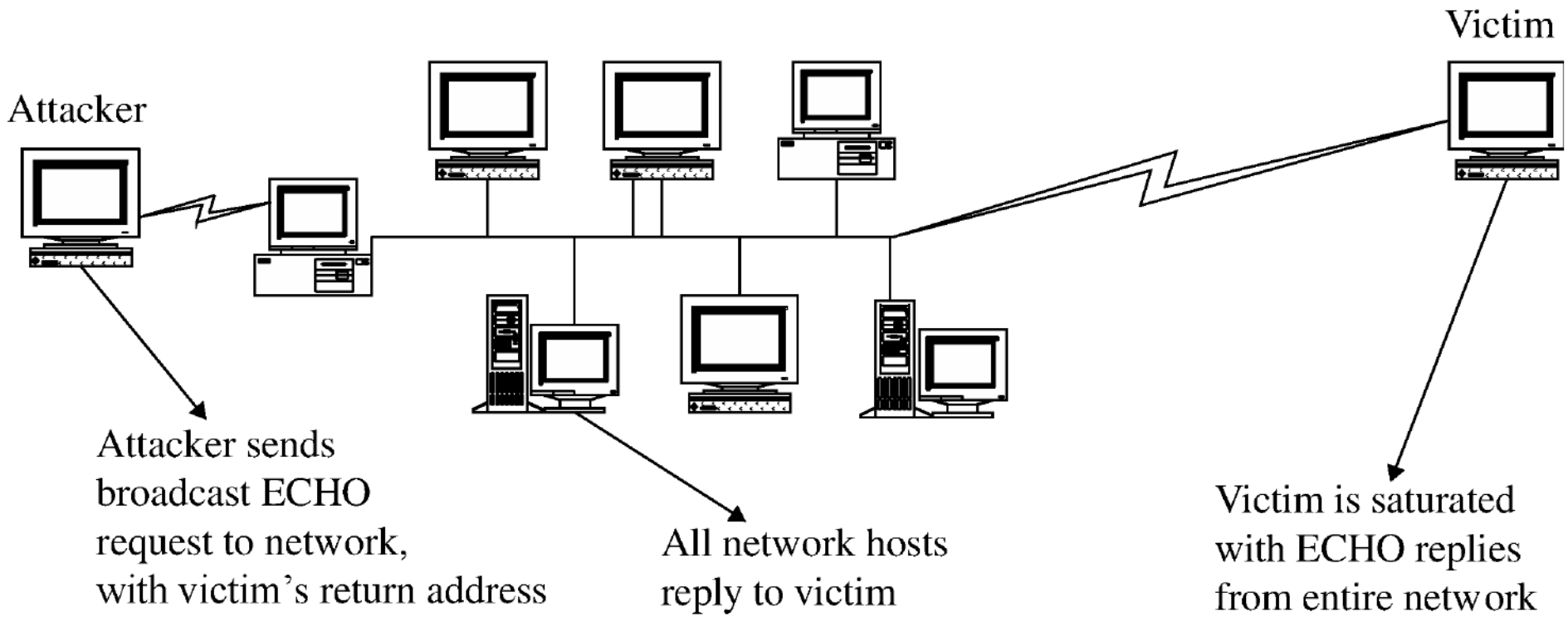


(a) Attacker has greater bandwidth



(b) Victim has greater bandwidth

DoS Attack: Smurf Attack



DoS Attack: Echo-Chargen

Victim A

Victim B

Chargen packet with echo bit on →

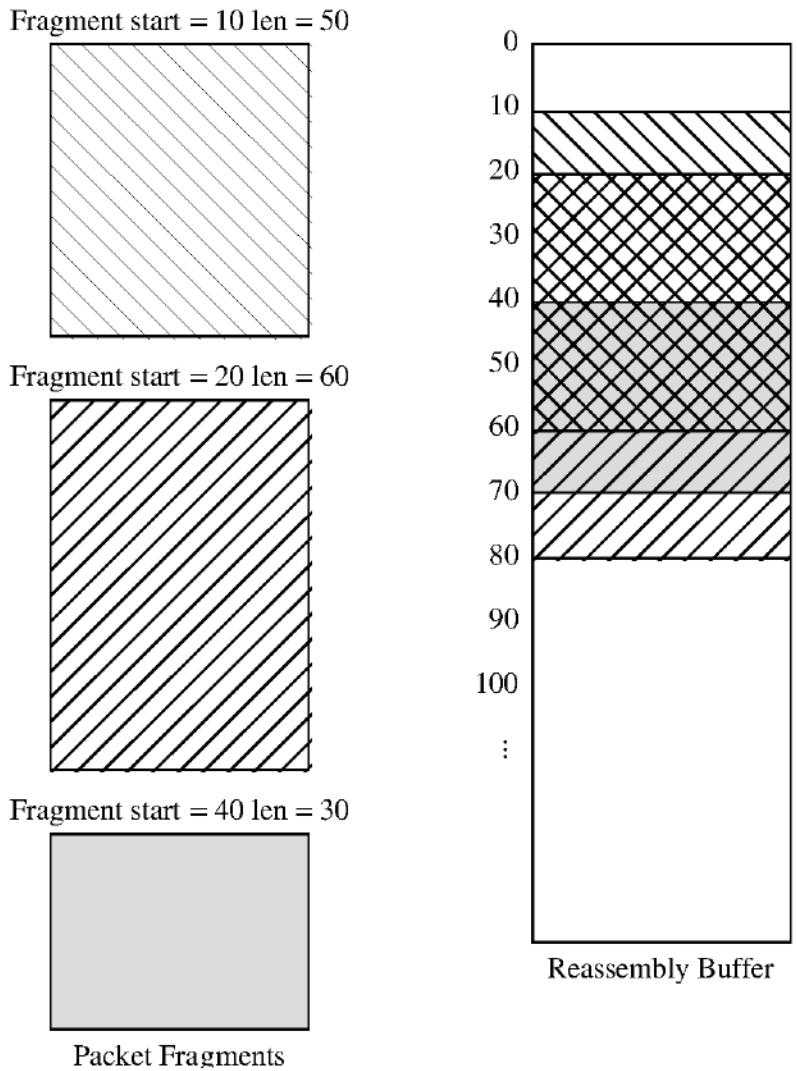
← Echoing what you just sent me

Chargen another packet with echo bit on →

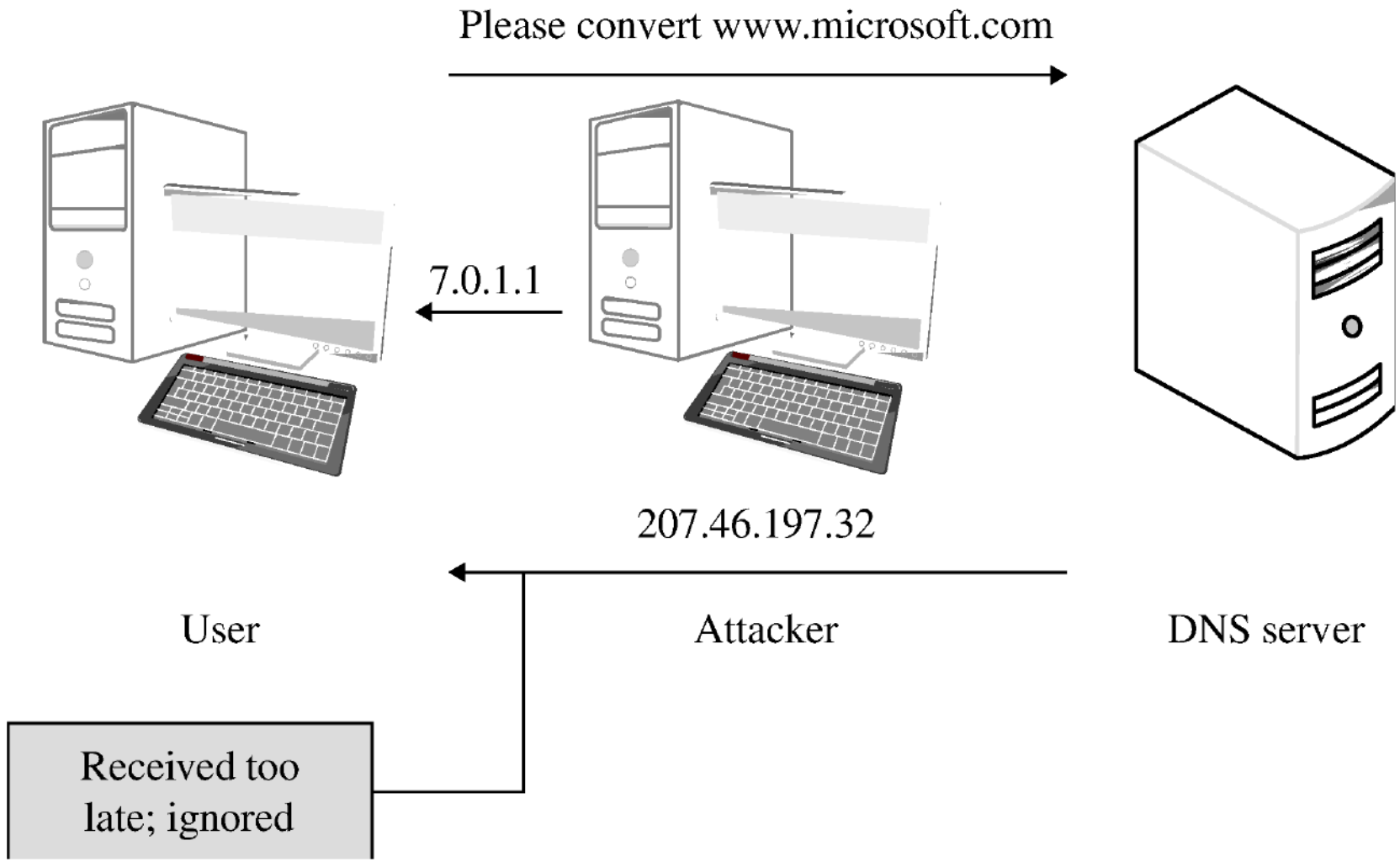
← Echoing that again

Chargen another packet with echo bit on →

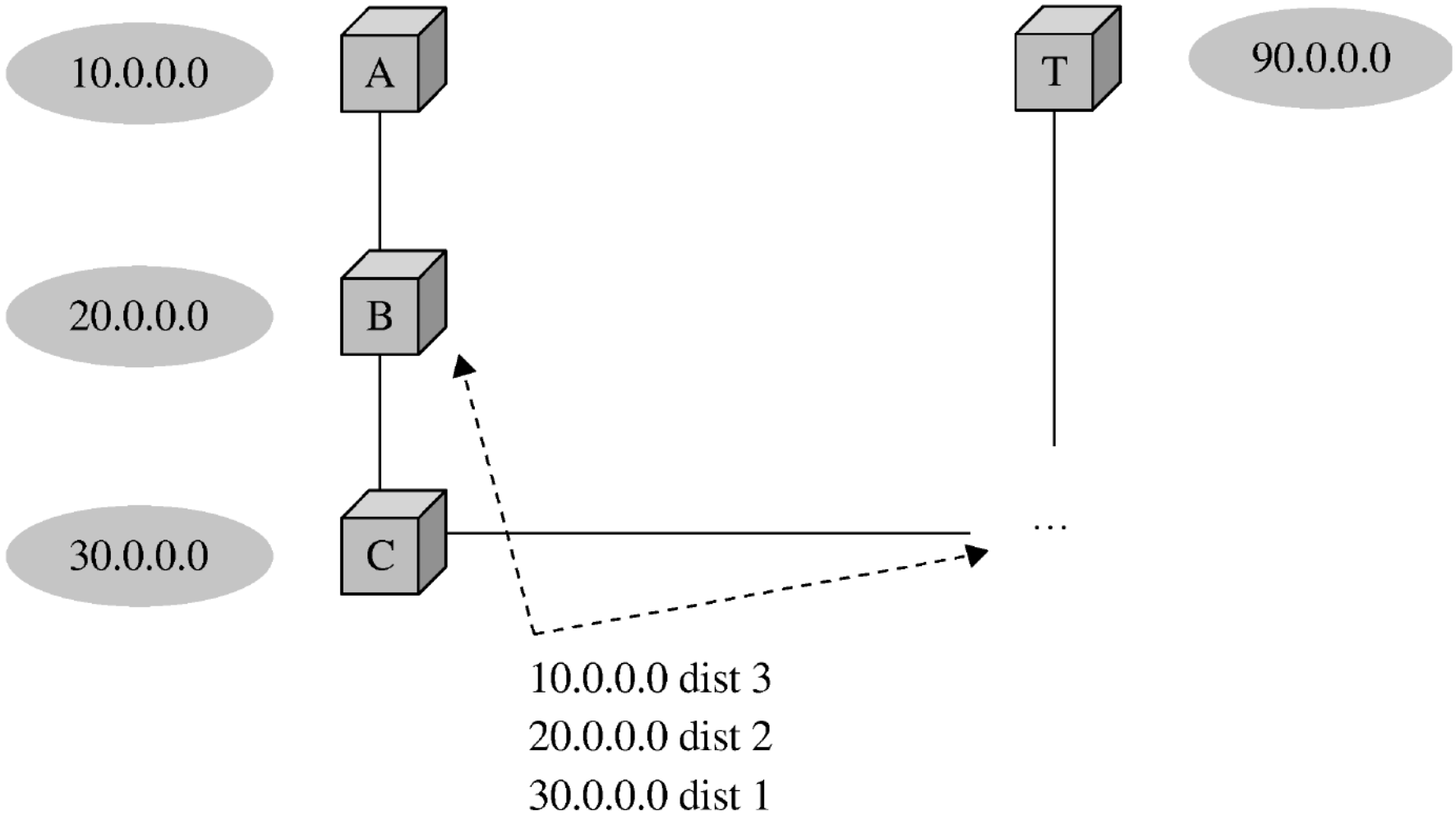
DoS Attack: Teardrop Attack



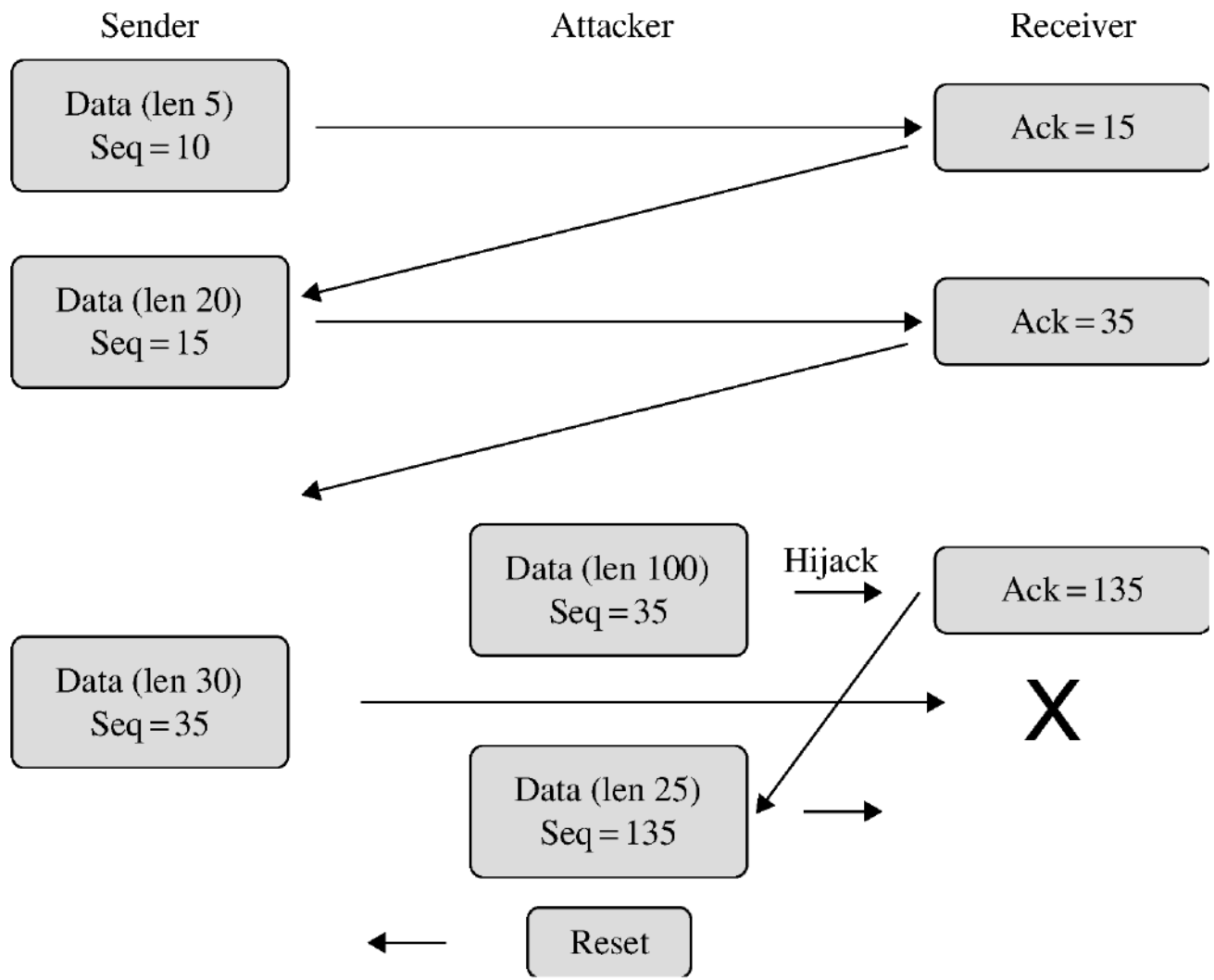
DoS Attack: DNS Spoofing



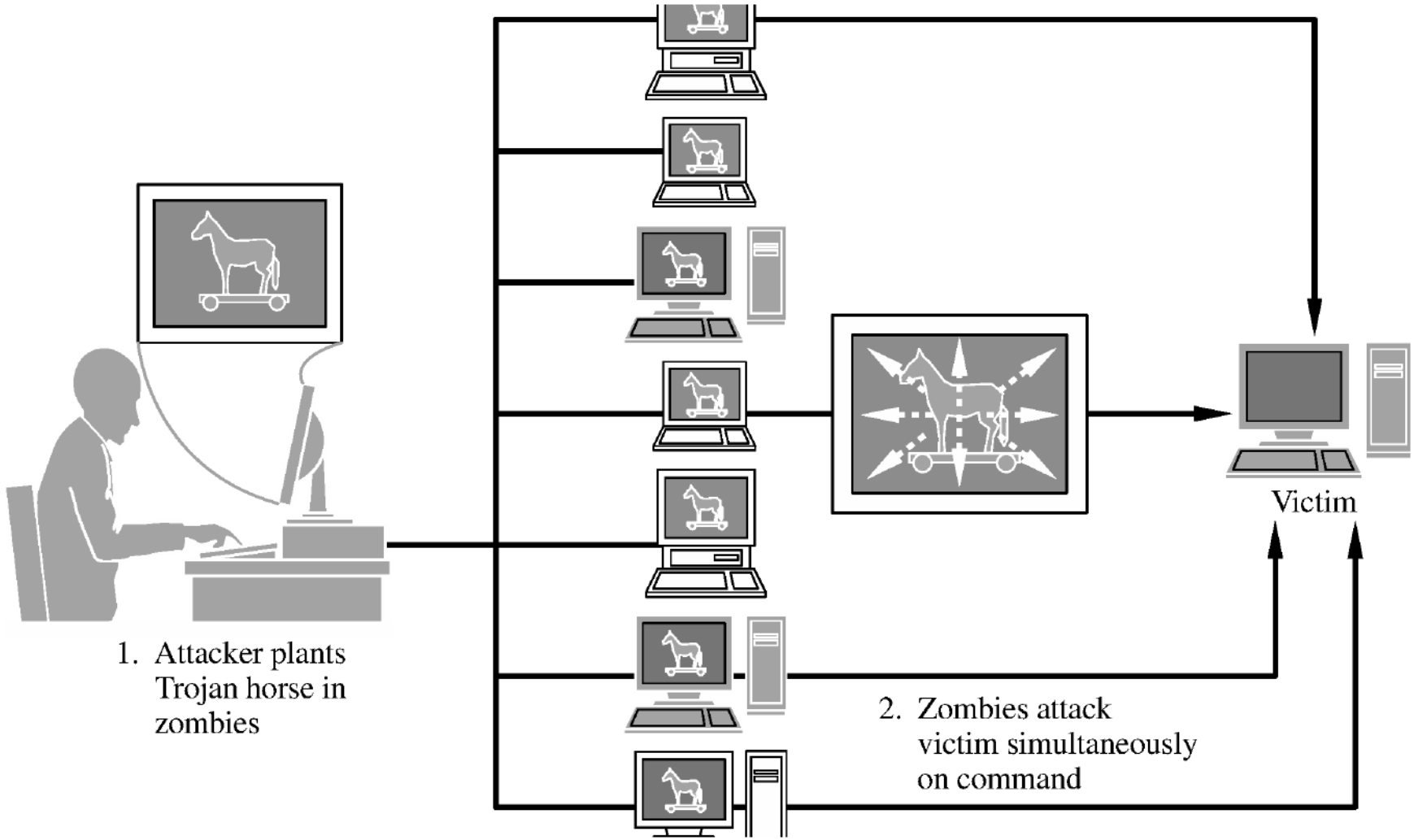
DoS Attack: Rerouting Routing



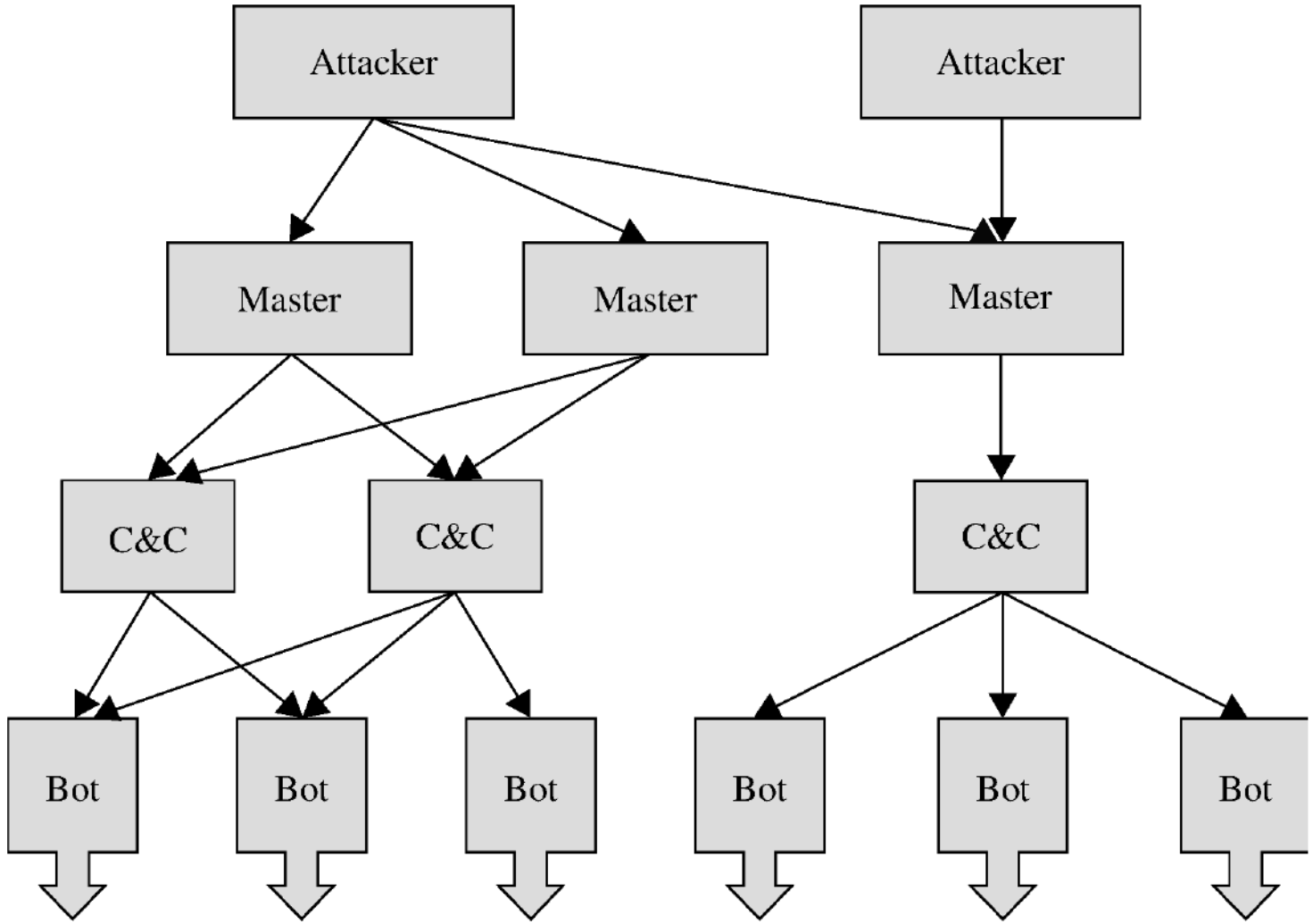
DoS Attack: Session Hijacking



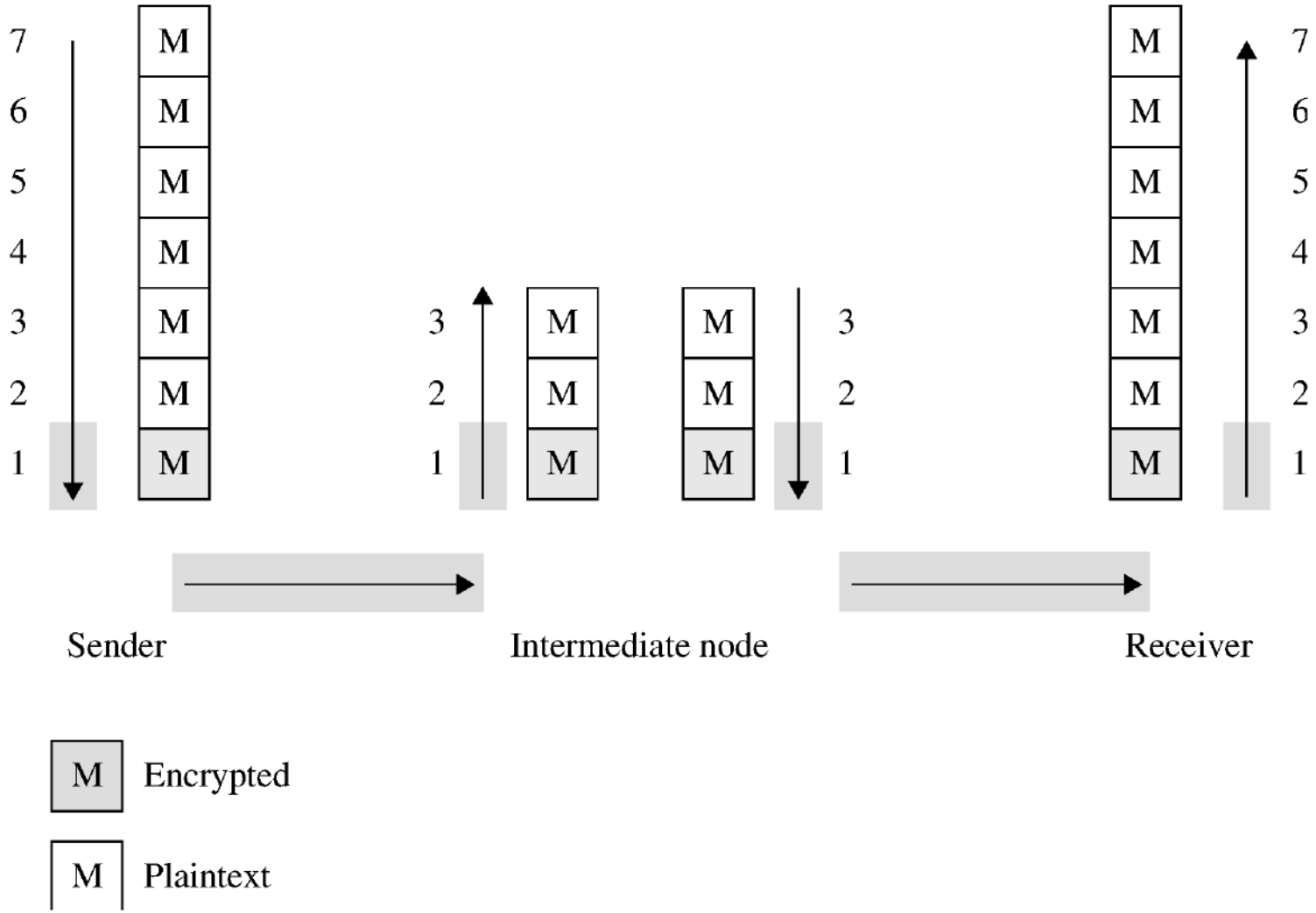
Distributed Denial of Service (DDoS)



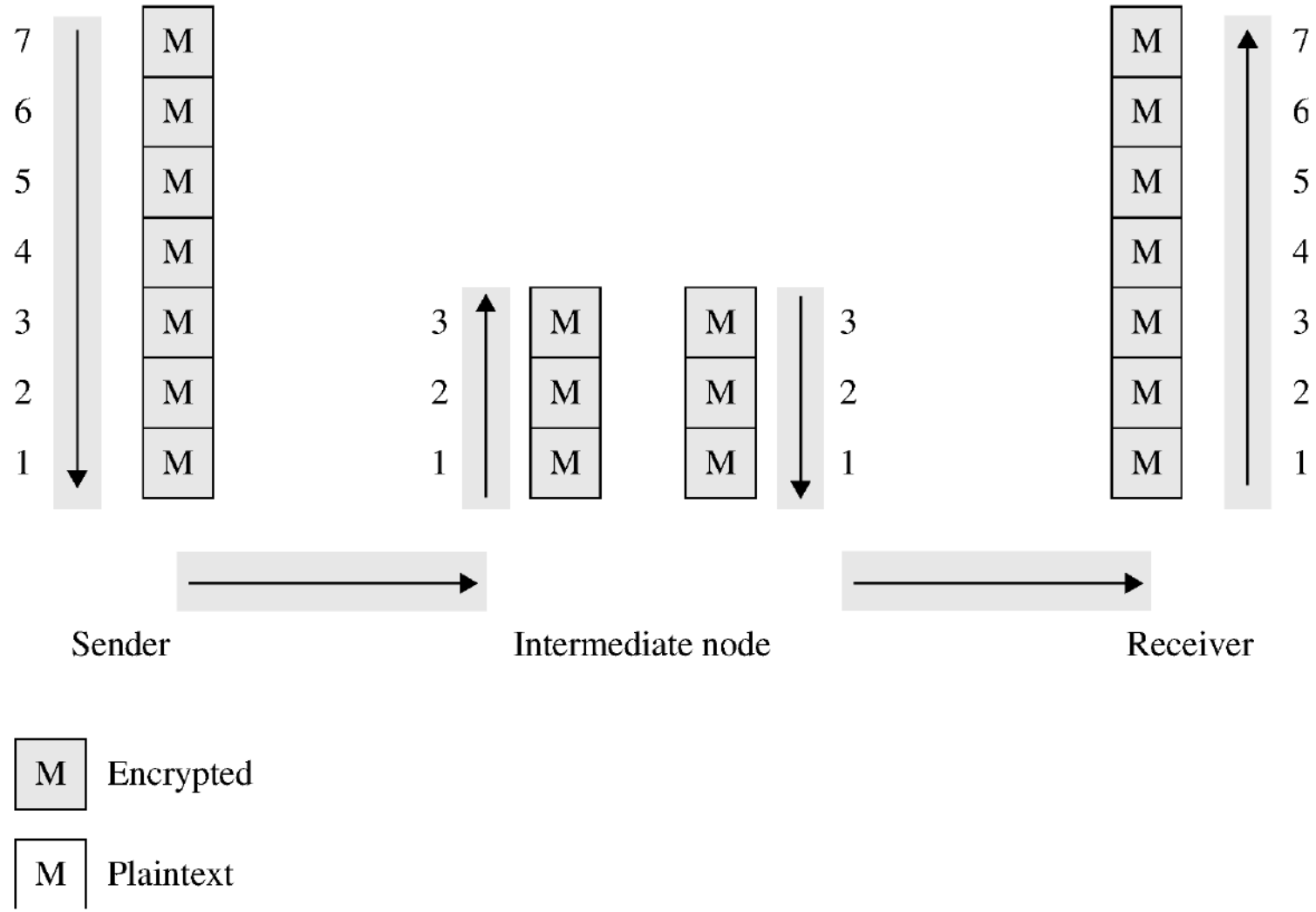
Botnets


Attacker
Attacker
Master
Master
Master
C&C
C&C
C&C
Bot
Bot
Bot
Bot
Bot
Bot

Link Encryption



End-to-End Encryption



Link vs. End-to-End

Link Encryption	End-to-End Encryption
Security within hosts	
Data partially exposed in sending host	Data protected in sending host
Data partially exposed in intermediate nodes	Data protected through intermediate nodes
Role of user	
Applied by sending host	Applied by user application
Invisible to user	User application encrypts
Host administrators select encryption	User selects algorithm
One facility for all users	Each user selects
Can be done in software or hardware	Usually software implementation; occasionally performed by user add-on hardware
All or no data encrypted	User can selectively encrypt individual data items
Implementation considerations	
Requires one key per pair of hosts	Requires one key per pair of users
Provides node authentication	Provides user authentication

Secure Shell (SSH)

- Originally developed for UNIX but now available on most OSs
- Provides an authenticated, encrypted path to the OS command line over the network
- Replacement for insecure utilities such as Telnet, rlogin, and rsh
- Protects against spoofing attacks and modification of data in communication

SSL and TLS

- Secure Sockets Layer (SSL) was designed in the 1990s to protect communication between a web browser and server
- In a 1999 upgrade to SSL, it was renamed Transport Layer Security (TLS)
- While the protocol is still commonly called SSL, TLS is the modern, and much more secure, protocol
- SSL is implemented at OSI layer 4 (transport) and provides
  - Server authentication
  - Client authentication (optional)
  - Encrypted communication

SSL Cipher Suites

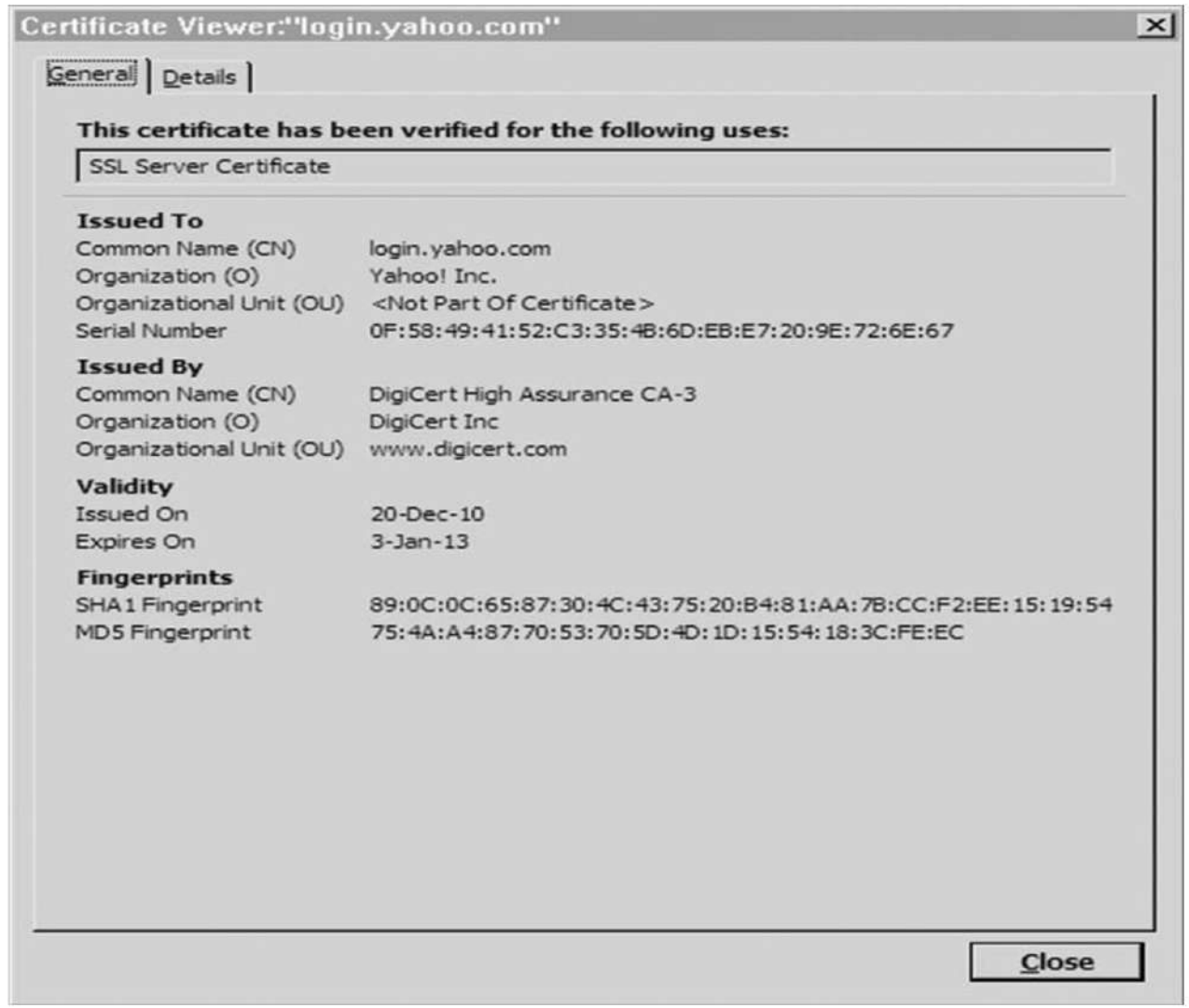
- At the start of an SSL session, the client and server negotiate encryption algorithms, known as the “cipher suite”
- The server sends a list of cipher suite options, and the client chooses an option from that list
- The cipher suite consists of
  - A digital signature algorithm for authentication
  - An encryption algorithm for confidentiality
  - A hash algorithm for integrity

SSL Cipher Suites (Partial List)

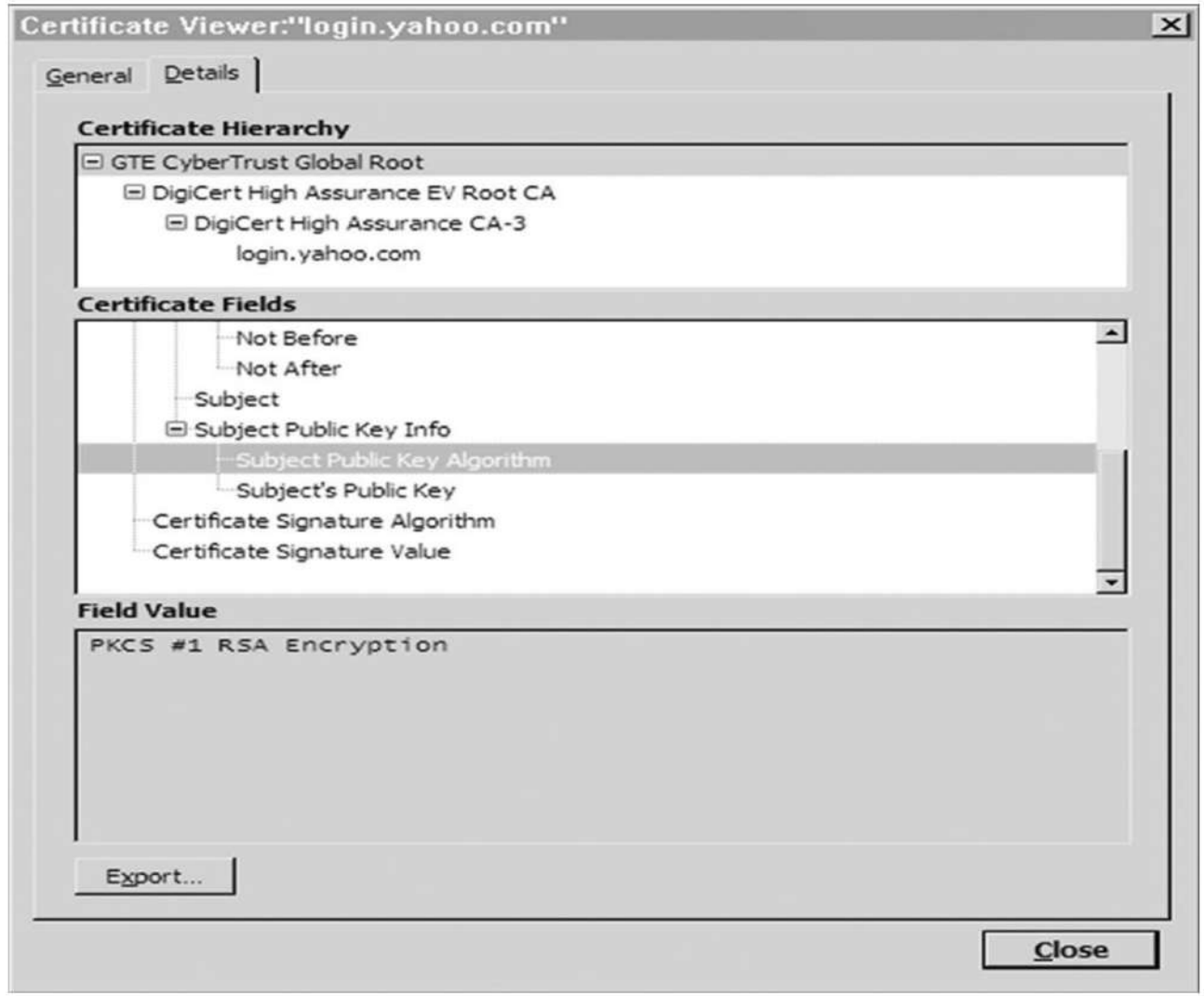
Cipher Suite Identifier	Algorithms Used
TLS_NULL_WITH_NULL_NULL	No authentication, no encryption, no hash function
TLS_RSA_WITH_NULL_MD5	RSA authentication, no encryption, MD5 hash function
TLS_RSA_EXPORT_WITH_RC4_40_MD5	RSA authentication with limited key length, RC4 encryption with a 40-bit key, MD5 hash function
TLS_RSA_WITH_3DES_EDE_CBC_SHA	RSA authentication, triple DES encryption, SHA-1 hash function
TLS_RSA_WITH_AES_128_CBC_SHA	RSA authentication, AES with a 128-bit key encryption, SHA-1 hash function
TLS_RSA_WITH_AES_256_CBC_SHA	RSA authentication, AES with a 256-bit key encryption, SHA-1 hash function
TLS_RSA_WITH_AES_128_CBC_SHA256	RSA authentication, AES with a 128-bit key encryption, SHA-256 hash function
TLS_RSA_WITH_AES_256_CBC_SHA256	RSA authentication, AES with a 256-bit key encryption, SHA-256 hash function
TLS_DH_DSS_WITH_3DES_EDE_CBC_SHA	Diffie-Hellman digital signature standard, triple DES encryption, SHA-1 hash function
TLS_RSA_WITH_CAMELLIA_256_CBC_SHA http://www.iana.org/go/rfc5932	RSA digital signature, Camellia encryption with a 256-bit key, SHA-1 hash function
TLS_ECDHE_ECDSA_WITH_ARIA_256_CBC_SHA384	Elliptic curve cryptosystem digital signature algorithm, Aria encryption with a 256-bit key, SHA-384 hash function

SSL Session Established

SSL Certificate



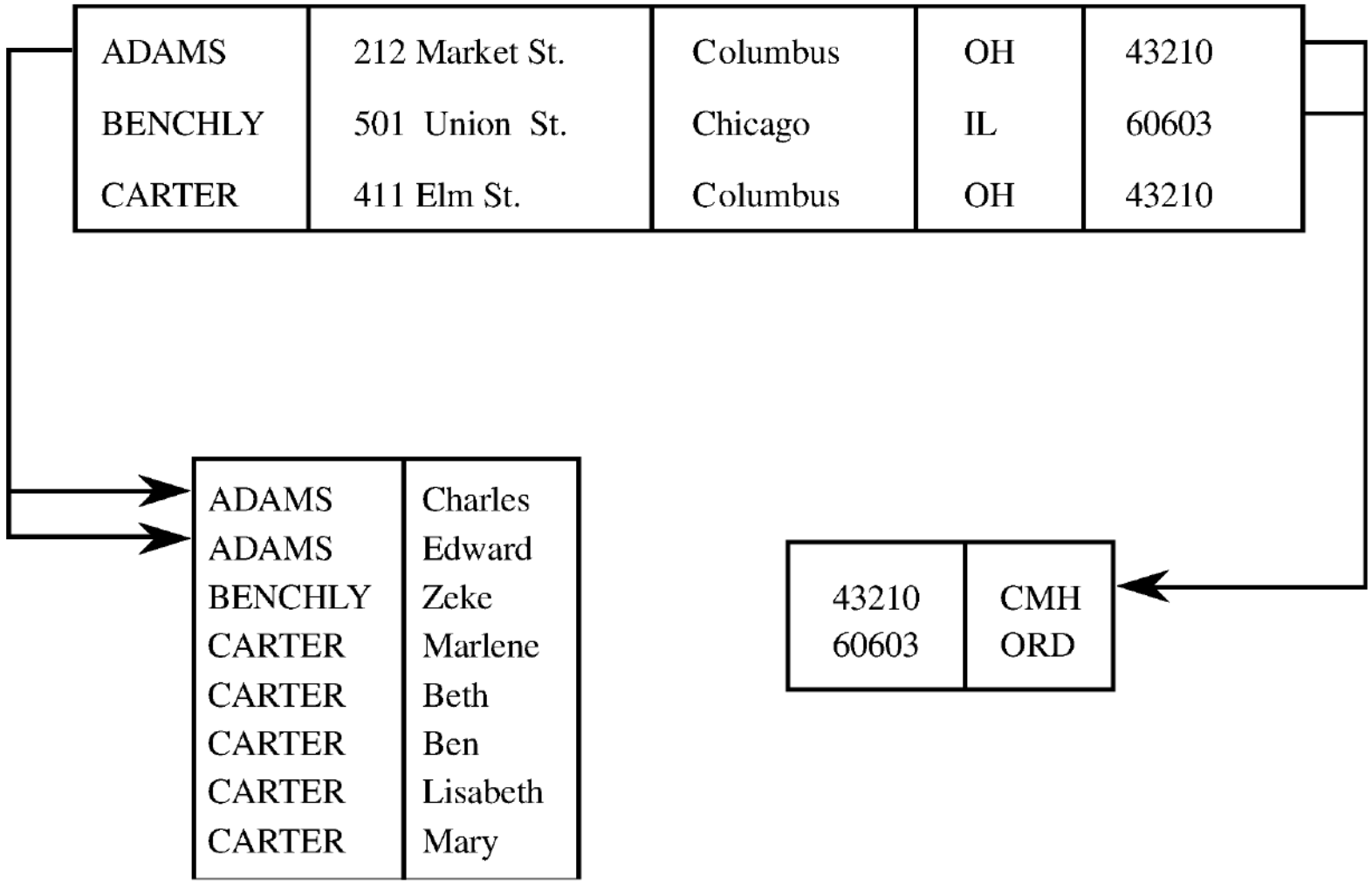
Chain of Certificates



Database Terms

- Database administrator
- Database management system (DBMS)
- Record
- Field/element
- Schema
- Subschema
- Attribute
- Relation

Database Example



Schema Example

Name	First	Address	City	State	Zip	Airport
ADAMS	Charles	212 Market St.	Columbus	OH	43210	CMH
ADAMS	Edward	212 Market St.	Columbus	OH	43210	CMH
BENCHLY	Zeke	501 Union St.	Chicago	IL	60603	ORD
CARTER	Marlene	411 Elm St.	Columbus	OH	43210	CMH
CARTER	Beth	411 Elm St.	Columbus	OH	43210	CMH
CARTER	Ben	411 Elm St.	Columbus	OH	43210	CMH
CARTER	Lisabeth	411 Elm St.	Columbus	OH	43210	CMH
CARTER	Mary	411 Elm St.	Columbus	OH	43210	CMH

Queries

- A query is a command that tells the database to retrieve, modify, add, or delete a field or record
- The most common database query language is SQL

Example SQL Query

- `SELECT ZIP='43210'`

Name	First	Address	City	State	Zip	Airport
ADAMS	Charles	212 Market St.	Columbus	OH	43210	CMH
ADAMS	Edward	212 Market St.	Columbus	OH	43210	CMH
CARTER	Marlene	411 Elm St.	Columbus	OH	43210	CMH
CARTER	Beth	411 Elm St.	Columbus	OH	43210	CMH
CARTER	Ben	411 Elm St.	Columbus	OH	43210	CMH
CARTER	Lisabeth	411 Elm St.	Columbus	OH	43210	CMH
CARTER	Mary	411 Elm St.	Columbus	OH	43210	CMH

Database Security Requirements

- Physical integrity
- Logical integrity
- Element integrity
- Auditability
- Access control
- User authentication
- Availability

Reliability and Integrity

- Reliability: in the context of databases, reliability is the ability to run for long periods without failing
- Database integrity: concern that the database as a whole is protected against damage
- Element integrity: concern that the value of a specific data element is written or changed only by authorized users
- Element accuracy: concern that only correct values are written into the elements of a database

Two-Phase Update

- Phase 1: Intent
  - DBMS does everything it can, other than making changes to the database, to prepare for the update
    - Collects records, opens files, locks out users, makes calculations
  - DBMS commits by writing a commit flag to the database
- Phase 2: Write
  - DBMS completes all write operations
  - DBMS removes the commit flag
- If the DBMS fails during either phase 1 or phase 2, it can be restarted and repeat that phase without causing harm

Other Database Security Concerns

- Error detection and correction codes to protect data integrity
- For recovery purposes, a database can maintain a change log, allowing it to repeat changes as necessary when recovering from failure
- Databases use locks and atomic operations to maintain consistency
  - Writes are treated as atomic operations
  - Records are locked during write so they cannot be read in a partially updated state

Sensitive Data

- Inherently sensitive
  - Passwords, locations of weapons
- From a sensitive source
  - Confidential informant
- Declared sensitive
  - Classified document, name of an anonymous donor
- Part of a sensitive attribute or record
  - Salary attribute in an employment database
- Sensitive in relation to previously disclosed information
  - An encrypted file combined with the password to open it

Types of Disclosures

- Exact data
- Bounds
- Negative result
- Existence
- Probable value
- Direct inference
- Inference by arithmetic
- Aggregation
- Hidden data attributes
  - File tags
  - Geotags

Direct Inference

- Inference is a way to infer or derive sensitive data from non sensitive data. It's a subtle vulnerability.

Name	Sex	Race	Aid	Fines	Drugs	Dorm
Adams	M	C	5000	45.	1	Holmes
Bailey	M	B	0	0.	0	Grey
Chin	F	A	3000	20.	0	West
Dewitt	M	B	1000	35.	3	Grey
Earhart	F	C	2000	95.	1	Holmes
Fein	F	C	1000	15.	0	West
Groff	M	C	4000	0.	3	West
Hill	F	B	5000	10.	2	Holmes
Koch	F	C	0	0.	1	West
Liu	F	A	0	10.	2	Grey
Majors	M	C	2000	0.	2	Grey

Direct Attack

- In a direct attack, a user tries to determine values of sensitive fields by seeking them directly with queries that yield few records.
- The most successful technique is to form a query so specific that it matches exactly one data item.

```
List NAME where
        (SEX=M ^ DRUGS=1)  v
        (SEX≠M ^ SEX≠F)  v
        (DORM=AYRES)
```

Inference by Arithmetic

Inference by Arithmetic

Another procedure, used by the U.S. Census Bureau and other organizations that gather sensitive data, is to release only statistics. The organizations suppress individual names, addresses, or other characteristics by which a single individual can be recognized. Only neutral statistics, such as count, sum, and mean, are released.

The indirect attack seeks to infer a final result based on one or more intermediate statistical results. But this approach requires work outside the database itself. In particular, a statistical attack seeks to use some apparently anonymous statistical measure to infer individual data. In the following sections, we present several examples of indirect attacks on databases that report statistics.

Sum

- This report reveals that no female living in Grey is receiving financial aid. This approach often allows us to determine a negative result.

	Holmes	Grey	West	Total
M	5000	3000	4000	12000
F	7000	0	4000	11000
Total	12000	3000	8000	23000

TABLE 7-8 Table Showing Negative Result

Count, mean and median

Sex	Holmes	Grey	West	Total
M	1	3	1	5
F	2	1	3	6
Total	3	4	4	11

TABLE 7-9 Inference from Count and Sum Results

Mean

The arithmetic **mean** (average) allows exact disclosure if the attacker can manipulate the subject population. As a trivial example, consider salary. Given the number of employees, the mean salary for a company and the mean salary of all employees except the president, it is easy to compute the president's salary.

Median

By a slightly more complicated process, we can determine an individual value from the **median**, the midpoint of an ordered list of values. The attack requires finding selections having one point of intersection that happens to be exactly in the middle, as shown in

Inference vs Aggregation

- Inference is difficult to control because it can occur from algebraic calculations beyond the scope of database management systems.
- Aggregation is nearly impossible for a dbms to control because combining the data can occur outside the system, even by multiple colluding users.

Aggregation

- Related to the inference problem is aggregation, which means building sensitive results from less sensitive inputs.
- You think of police investigation- starting with entire population- narrowing the analysis – single person.
- But- if police officers- work – in parallel-
- list of possible suspects –
- another may have a list with possible motive –
- another may have list of capable person.
- When the intersection of these lists is single person, the police have their prime suspect.

Hidden data Attributes

- Objects such as pictures, music files, and documents are actually complex data structures having properties or attributes that add meaning to the data. These properties called meta data, are not displayed with the picture or document, but they are not concealed; infact numerous applications support selecting , searching , sorting and editing based on metadata.

File tags

- One use of attributes is tags for pictures.
- Tag for documents. Each document has properties that include the name of author, author’s organization, data created and data last saved etc.

Geo tags

- Adam savage – host of program mythbusters- took a photo of his car- front of his house. (Geotagging problem)
- Friedland and Sommer who studied the problem of geotagging, - between 1 to 5 percent of photos at sites such as flickr, youtube, and craiglist contain header data that gives the location where the pictures were taken.

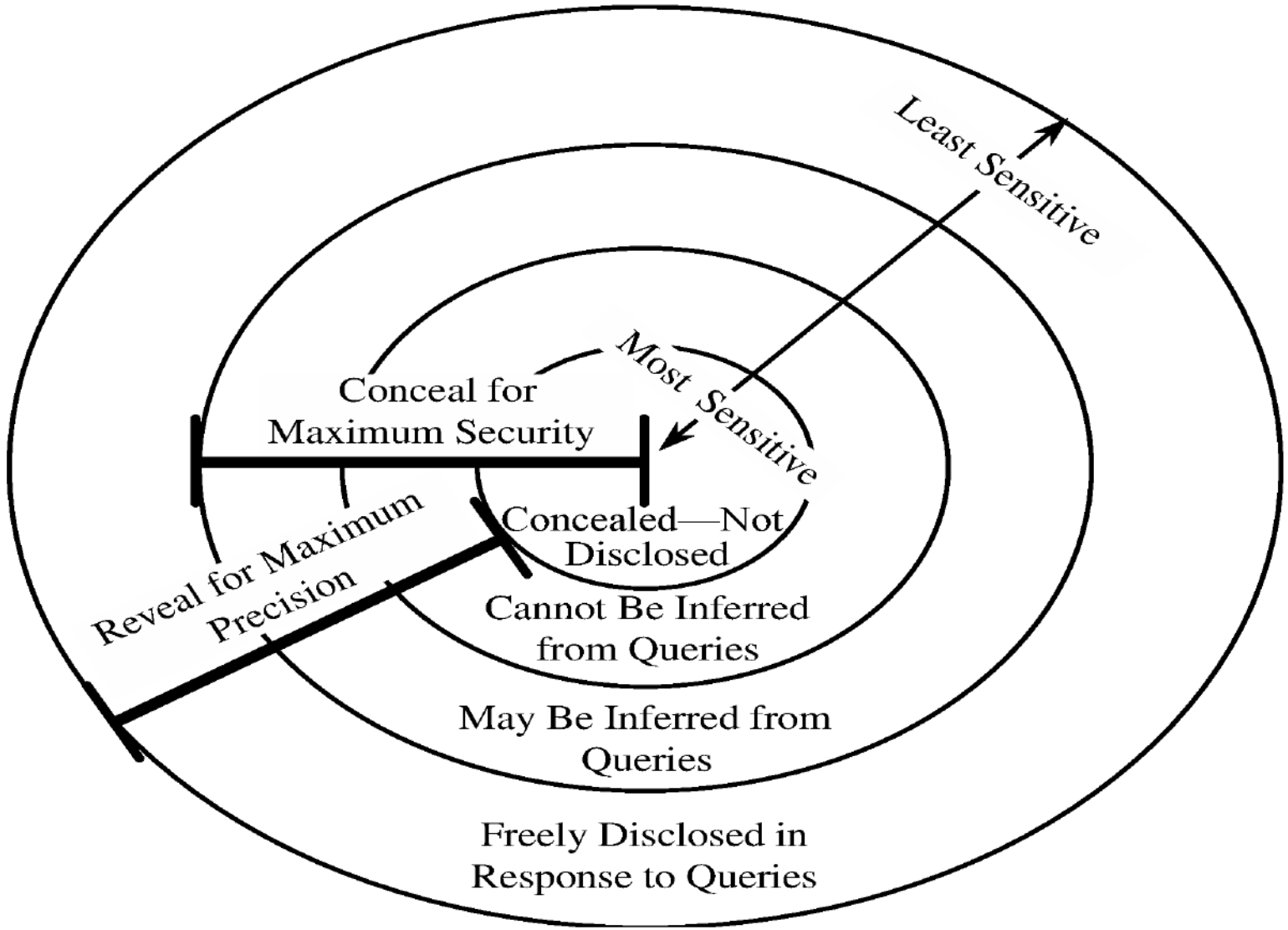
Preventing Disclosure: (Data Suppression and modification)

- There are no perfect solutions to inference and aggregation problems.
- 3 approaches to control them are:
- First two methods- used- to limit queries accepted or to limit data provided in response to a query.
- Third method applies only to data released.

1. Suppress obviously sensitive information
2. Keep **track** of **what** each **user knows** based on past queries
3. Disguise the data

- *Data Suppression blocks release of sensitive data*
- *Data concealing releases part or approximation of sensitive data*

Security vs. Precision



Suppression Techniques

- Limited response suppression
  - Eliminates certain low-frequency elements from being displayed
- Combined results
  - Ranges, rounding, sums, averages
- Random sample
- Blocking small sample sizes
- Random data perturbation
  - Randomly add or subtract a small error value to/from actual values
- Swapping
  - Randomly swapping values for individual records while keeping statistical results the same

Data Mining

- Data mining uses statistics, machine learning, mathematical models, pattern recognition, and other techniques to discover patterns and relations on large datasets
- The size and value of the datasets present an important security and privacy challenge, as the consequences of disclosure are naturally high

Data Mining Challenges

- Correcting mistakes in data
- Using Comparable data
- Eliminating false matches
- Availability of Data

Preserving privacy

Granular access control

Secure data storage

Transaction logs

Real-time security monitoring

Summary

- Database security requirements include:
  - Physical integrity
  - Logical integrity
  - Element integrity
  - Auditability
  - Access control
  - User authentication
  - Availability
- There are many subtle ways for sensitive data to be inadvertently disclosed, and there is no single answer for prevention
- Data mining and big data have numerous open security and privacy challenges

Case 3: Confidentiality + Authentication