

# COM307000 - Access Control Authorization

Dr. Anca Jurcut

E-mail: [anca.jurcut@ucd.ie](mailto:anca.jurcut@ucd.ie)

School of Computer Science and Informatics  
University College Dublin,  
Ireland



# Authorization

# Authentication vs Authorization

- ❑ **Authentication** — Are you who you say you are?
  - Restrictions on who (or what) can access system
- ❑ **Authorization** — Are you allowed to do that?
  - Restrictions on actions of authenticated users
- ❑ Authorization is a form of **access control**
- ❑ But first, we look at system certification...

# System Certification

- ❑ Government attempt to certify “security level” of products
- ❑ Of historical interest
  - Sorta like a history of authorization
- ❑ Still important today if you want to sell a product to the government
  - Tempting to argue it's a failure since government is so insecure, but...

# Orange Book

- ❑ Trusted Computing System Evaluation Criteria (TCSEC), 1983
  - Universally known as the “orange book”
  - Name is due to color of it's cover
  - About 115 pages
  - Developed by U.S. DoD (NSA)
  - Part of the “rainbow series”
- ❑ Orange book generated a pseudo-religious fervor among some people
  - Less and less intensity as time goes by

# Orange Book Outline

- Goals

- Provide way to assess security products
- Provide general guidance/philosophy on how to build more secure products

- Four ***divisions*** labeled D thru A

- D is lowest, A is highest

- Divisions split into numbered ***classes***

# EAL 1 thru 7

- ❑ EAL1 — functionally tested
- ❑ EAL2 — structurally tested
- ❑ EAL3 — methodically tested, checked
- ❑ EAL4 — ***designed***, tested, reviewed
- ❑ EAL5 — semiformally designed, tested
- ❑ EAL6 — verified, designed, tested
- ❑ EAL7 — formally ... (blah blah blah)

# Authentication vs Authorization

- ❑ Authentication — Are you who you say you are?
  - Restrictions on who (or what) can access system
- ❑ **Authorization** — Are you allowed to do that?
  - Restrictions on actions of authenticated users
- ❑ Authorization is a form of **access control**
- ❑ Classic view of authorization...
  - Access Control Lists (ACLs)
  - Capabilities (C-lists)



# Lampson's Access Control Matrix

- ❑ **Subjects** (users) index the rows
- ❑ **Objects** (resources) index the columns

	OS	Accounting program	Accounting data	Insurance data	Payroll data
Bob	rx	rx	r	—	—
Alice	rx	rx	r	rw	rw
Sam	rwX	rwX	r	rw	rw
Accounting program	rx	rx	rw	rw	rw

# Are You Allowed to Do That?

- ❑ **Access control matrix** has **all** relevant info
- ❑ Could be 100's of users, 10,000's of resources
  - Then matrix with 1,000,000's of entries
- ❑ How to manage such a large matrix?
- ❑ Note: We need to check this matrix before access to any resource by any user
- ❑ How to make this efficient/practical?

# Access Control Lists (ACLs)

- ❑ ACL: store access control matrix by **column**
- ❑ Example: ACL for **insurance data** is in **blue**

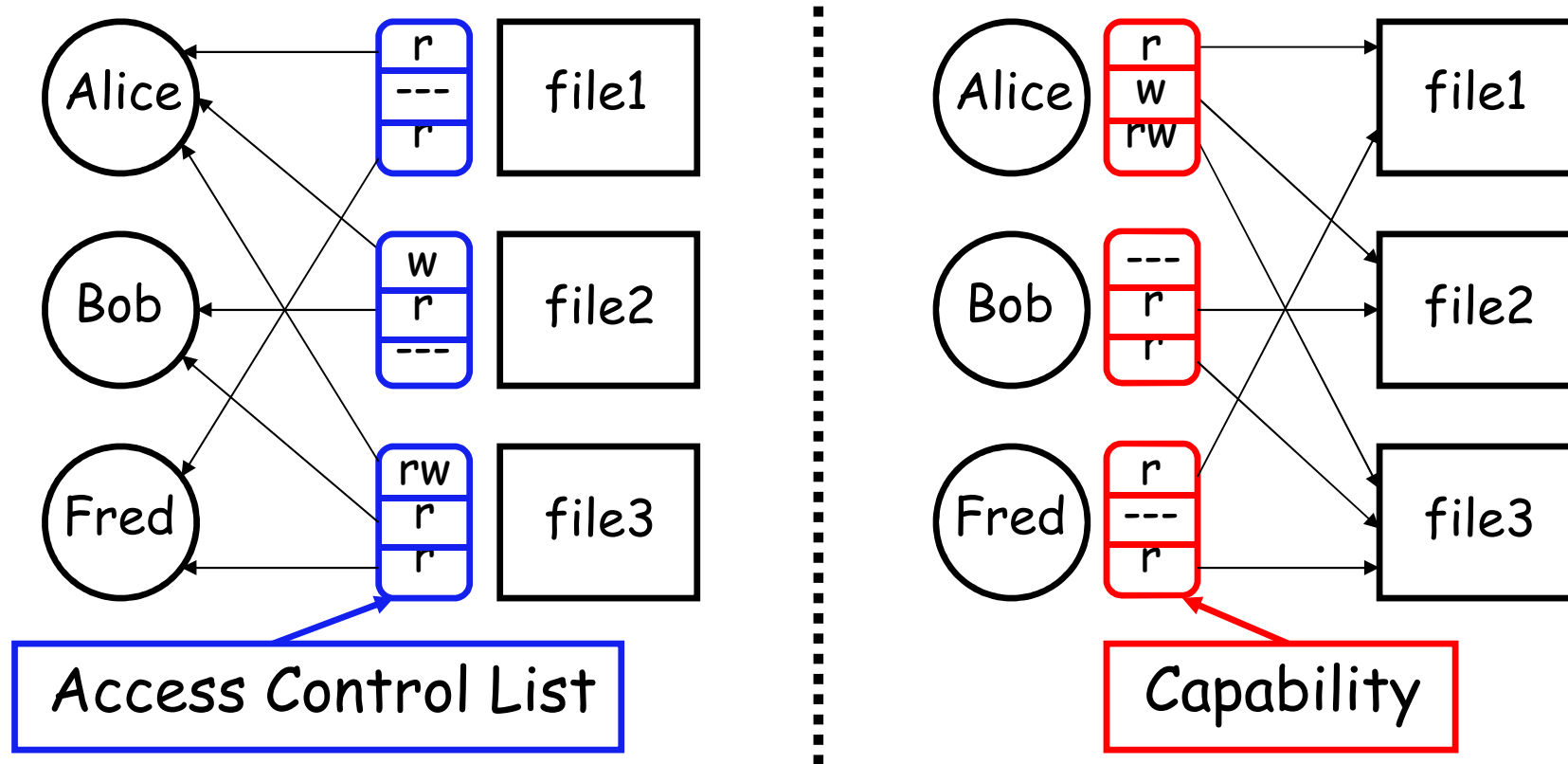
	OS	Accounting program	Accounting data	<b>Insurance data</b>	Payroll data
Bob	rx	rx	r	—	—
Alice	rx	rx	r	<b>rw</b>	rw
Sam	rwX	rwX	r	<b>rw</b>	rw
Accounting program	rx	rx	rw	<b>rw</b>	rw

# Capabilities (or C-Lists)

- ❑ Store access control matrix by **row**
- ❑ Example: Capability for **Alice** is in **red**

	OS	Accounting program	Accounting data	Insurance data	Payroll data
Bob	rx	rx	r	—	—
Alice	rx	rx	r	rw	rw
Sam	rwX	rwX	r	rw	rw
Accounting program	rx	rx	rw	rw	rw

# ACLs vs Capabilities



- ❑ Note that arrows point in opposite directions...
- ❑ With ACLs, still need to associate users to files

# ACLs vs Capabilities

## ❑ ACLs

- Good when users manage their own files
- Protection is data-oriented
- Easy to change rights to a resource

## ❑ Capabilities

- Easy to delegate — avoid the [confused deputy](#)
- Easy to add/delete users
- More difficult to implement
- The “Zen of information security”

## ❑ Capabilities loved by academics

- [Capability Myths Demolished](#)

# Multilevel Security (MLS) Models

# Classifications and Clearances

- ❑ **Classifications** apply to **objects**
- ❑ **Clearances** apply to **subjects**
- ❑ US Department of Defense (DoD) uses 4 levels:

**TOP SECRET**

**SECRET**

**CONFIDENTIAL**

**UNCLASSIFIED**



# Multilevel Security (MLS)

- ❑ MLS needed when subjects/objects at different levels access **same system**
- ❑ MLS is a form of **Access Control**
- ❑ Military and government interest in MLS for many decades
  - Lots of research into MLS
  - Strengths and weaknesses of MLS well understood (almost entirely theoretical)
  - Many possible uses of MLS outside military

# MLS Applications

- ❑ Classified government/military systems
- ❑ **Business example:** info restricted to
  - Senior management only, all management, everyone in company, or general public
- ❑ Network firewall
- ❑ Confidential medical info, databases, etc.
- ❑ Usually, MLS not really a technical system
  - More like part of a legal structure

# MLS Security Models

- ❑ MLS models explain **what** needs to be done
- ❑ Models **do not** tell you **how** to implement
- ❑ Models are descriptive, not prescriptive
  - That is, high-level description, not an algorithm
- ❑ There are many MLS models
- ❑ We'll discuss simplest MLS model
  - Other models are more realistic
  - Other models also more complex, more difficult to enforce, harder to verify, etc.

# Bell-LaPadula

- ❑ BLP security model designed to express essential requirements for MLS
- ❑ BLP deals with **confidentiality**
  - To prevent unauthorized reading
- ❑ Recall that  $O$  is an object,  $S$  a subject
  - Object  $O$  has a classification
  - Subject  $S$  has a clearance
  - Security level denoted  $L(O)$  and  $L(S)$

# BLP: The Bottom Line

- ❑ BLP is simple, probably too simple
- ❑ BLP is one of the few security models that can be used to prove things about systems
- ❑ BLP has inspired other security models
  - Most other models try to be more realistic
  - Other security models are more complex
  - Models difficult to analyze, apply in practice

# Biba's Model

- ❑ BLP for confidentiality, Biba for **integrity**
  - Biba is to prevent unauthorized writing
- ❑ Biba is (in a sense) the dual of BLP
- ❑ Integrity model
  - Spse you trust the integrity of **○** but not **○**
  - If object **○** includes **○** and **○** then you cannot trust the integrity of **○**
- ❑ Integrity level of **○** is minimum of the integrity of any object in **○**
- ❑ **Low water mark** principle for integrity

# Compartments

# Compartments

- ❑ Multilevel Security (MLS) enforces access control **up and down**
- ❑ Simple hierarchy of security labels is generally *not* flexible enough
- ❑ Compartments enforces restrictions **across**
- ❑ Suppose **TOP SECRET** divided into **TOP SECRET {CAT}** and **TOP SECRET {DOG}**
- ❑ Both are **TOP SECRET** but information flow restricted across the **TOP SECRET** level

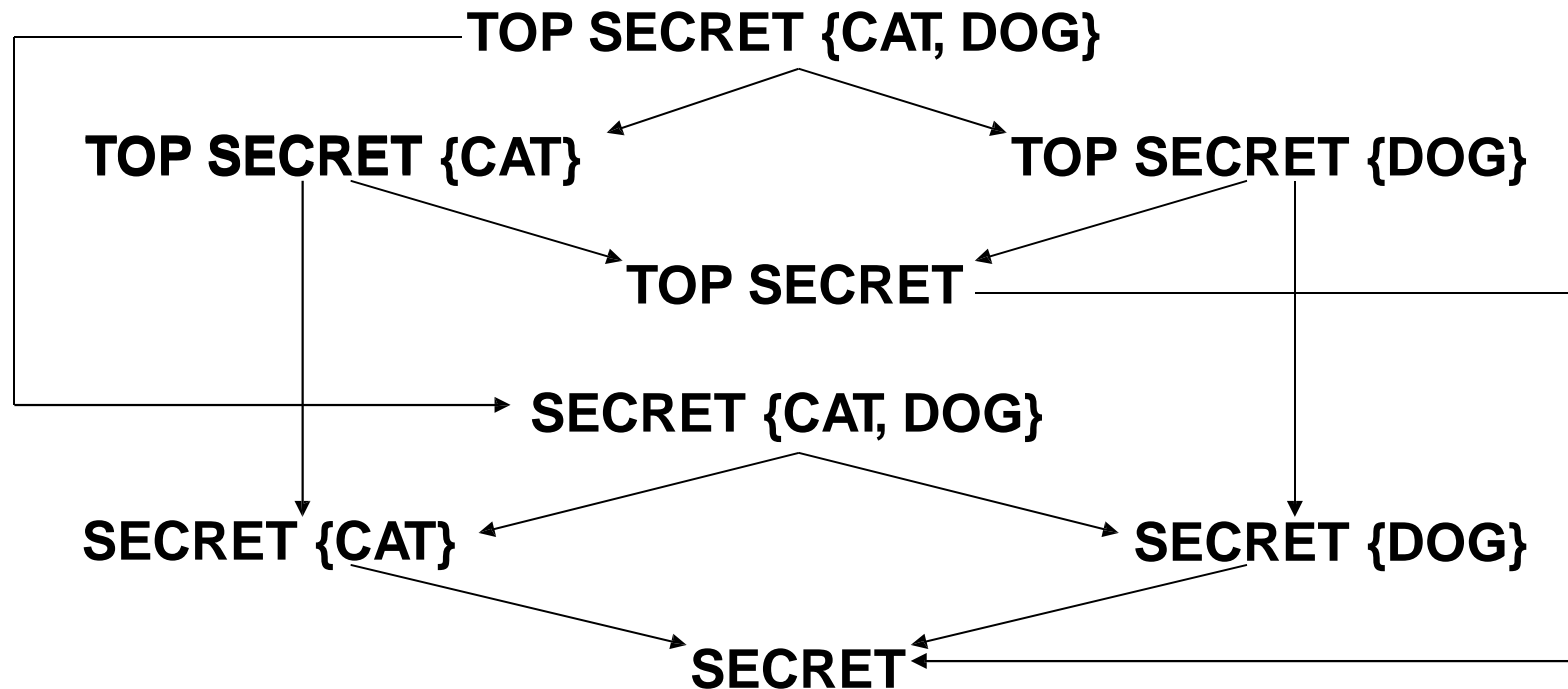


# Compartments

- ❑ Why compartments?
  - Why not create a new classification level?
- ❑ May ***not*** want either of
  - **TOP SECRET {CAT}  $\geq$  TOP SECRET {DOG}**
  - **TOP SECRET {DOG}  $\geq$  TOP SECRET {CAT}**
- ❑ Compartments designed to enforce the **need to know** principle
  - Regardless of clearance, you only have access to info that you need to know to do your job

# Compartments

- Arrows indicate “ $\geq$ ” relationship



- Not all classifications are comparable, e.g.,  
**TOP SECRET {CAT}** vs **SECRET {CAT, DOG}**

# MLS vs Compartments

- ❑ MLS can be used without compartments
  - And vice-versa
- ❑ But, MLS almost always uses compartments
- ❑ Example
  - MLS mandated for protecting medical records of British Medical Association (BMA)
  - AIDS was **TOP SECRET**, prescriptions **SECRET**
  - Everything tends toward **TOP SECRET**
  - Defeats the purpose of the system!
  - Compartments-only approach used instead

# Covert Channel

# Covert Channel

- ❑ MLS designed to restrict legitimate channels of communication
- ❑ May be other ways for information to flow
- ❑ For example, resources shared at different levels could be used to “signal” information
- ❑ **Covert channel**: a communication path not intended as such by system’s designers

# Covert Channel Example

- ❑ Alice has **TOP SECRET** clearance, Bob has **CONFIDENTIAL** clearance
- ❑ Suppose the file space shared by all users
- ❑ Alice creates file FileXYzW to signal “1” to Bob, and removes file to signal “0”
- ❑ Once per minute Bob lists the files
  - If file FileXYzW does not exist, Alice sent 0
  - If file FileXYzW exists, Alice sent 1
- ❑ Alice can leak **TOP SECRET** info to Bob

# Covert Channel Example

**Alice:**    Create file    Delete file    Create file                      Delete file

**Bob:**    Check file    Check file    Check file    Check file    Check file

**Data:**                1                0                1                1                0

**Time:**    

# Covert Channel

- ❑ Other possible covert channels?
  - Print queue
  - ACK messages
  - Network traffic, etc.
- ❑ When does covert channel exist?
  1. Sender and receiver have a shared resource
  2. Sender able to vary some property of resource that receiver can observe
  3. “Communication” between sender and receiver can be synchronized

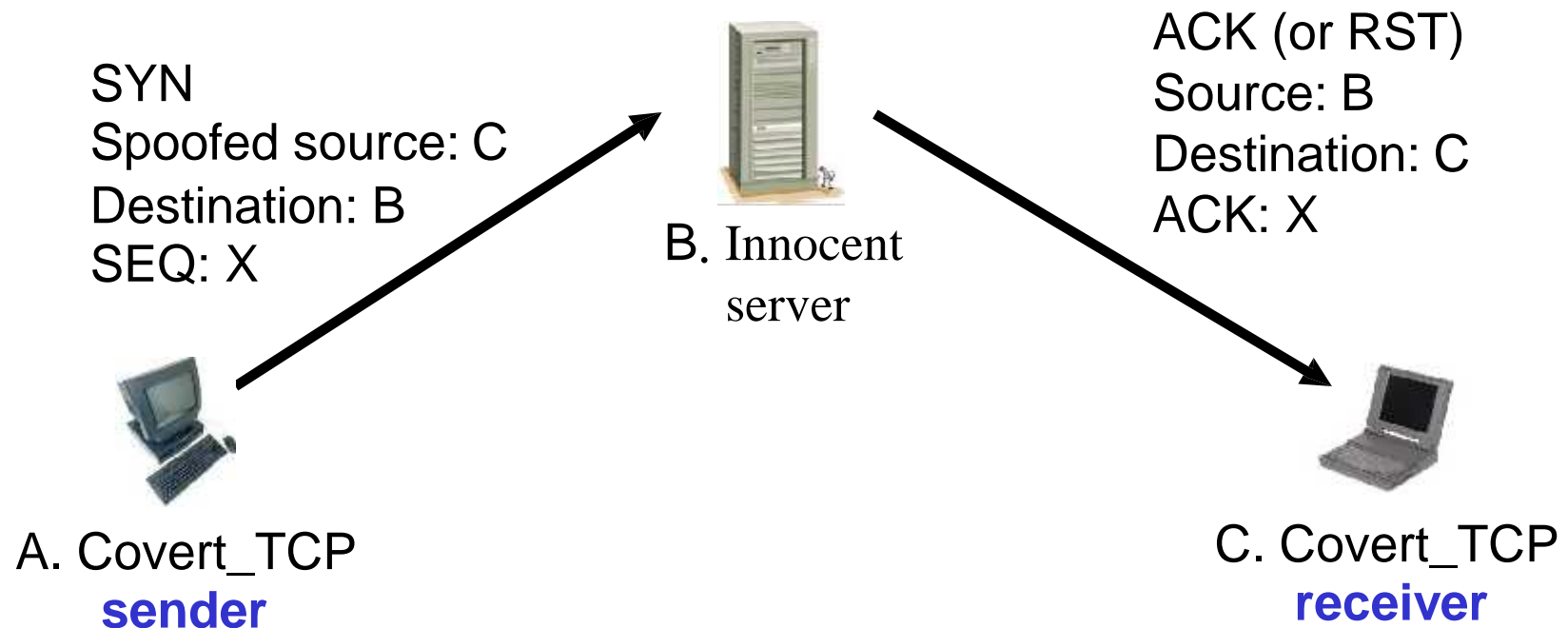


# Covert Channel

- ❑ Potential covert channels are everywhere
- ❑ But, it's easy to eliminate covert channels:
  - “Just” eliminate all shared resources and all communication!
- ❑ Virtually impossible to eliminate covert channels in any **useful** information system
  - DoD guidelines: **reduce covert channel capacity** to no more than 1 bit/second
  - Implication? DoD has given up on *eliminating* covert channels

# Real-World Covert Channel

- ❑ Hide data in TCP sequence numbers
- ❑ Tool: covert\_TCP
- ❑ Sequence number X contains covert info



# Inference Control

# Inference Control Example

- ❑ Suppose we query a database
  - Question: What is average salary of female CS professors at UCD?
  - Answer: 90,000 euros
  - Question: How many female CS professors at UCD?
  - Answer: 1
- ❑ Specific information has leaked from responses to general questions!

# Inference Control & Research

- ❑ For example, medical records are private but valuable for research
- ❑ How to make info available for research and protect privacy?
- ❑ How to allow access to such data without leaking specific information?

# Naiïve Inference Control

- ❑ Remove names from medical records?
- ❑ Still may be easy to get specific info from such “anonymous” data
- ❑ Removing names is not enough
  - As seen in previous example
- ❑ What more can be done?

# Less-naïve Inference Control

- ❑ Query set size control
  - Don't return an answer if set size is too small
- ❑ N-respondent, k% dominance rule
  - Do not release statistic if k% or more contributed by N or fewer
  - Example: Avg salary in Bill Gates' neighborhood
  - This approach used by US Census Bureau
- ❑ Randomization
  - Add small amount of random noise to data
- ❑ Many other methods — none satisfactory

# Netflix Example

- ❑ Netflix prize — \$1M to first to improve recommendation system by 10% or more
- ❑ Netflix created dataset for contest
  - Movie preferences of real users
  - Usernames removed, some “noise” added
- ❑ Insufficient inference control
  - Researchers able to correlate IMDB reviews with those in Netflix dataset



# Something Better Than Nothing?

- ❑ Robust inference control may be impossible
- ❑ Is weak inference control better than nothing?
  - **Yes**: Reduces amount of information that leaks
- ❑ Is weak covert channel protection better than nothing?
  - **Yes**: Reduces amount of information that leaks
- ❑ Is weak crypto better than no crypto?
  - **Probably not**: Encryption indicates important data
  - May be easier to filter encrypted data

CAPTCHA

# Turing Test

- ❑ Proposed by Alan Turing in 1950
- ❑ Human asks questions to a human and a computer, without seeing either
- ❑ If questioner cannot distinguish human from computer, computer passes
- ❑ This is the **gold standard** in AI
- ❑ No computer can pass this today
  - But some claim they are close to passing

# CAPTCHA

## ❑ CAPTCHA

- Completely Automated Public Turing test to tell Computers and Humans Apart

## ❑ Completely Automated — test is generated and scored by a computer

## ❑ Public — program and data are public

## ❑ Turing test to tell... — humans can pass the test, but machines cannot

- Also known as HIP == Human Interactive Proof

## ❑ Like an inverse Turing test (sort of...)

# CAPTCHA Paradox?

- ❑ “...CAPTCHA is a program that can generate and grade tests that it itself cannot pass...”
- ❑ “...much like some professors...”
- ❑ Paradox — computer creates and scores test that it itself cannot pass!
- ❑ CAPTCHA purpose?
  - Only humans get access (not bots/computers)
- ❑ So, CAPTCHA is for **access control**

# CAPTCHA Uses?

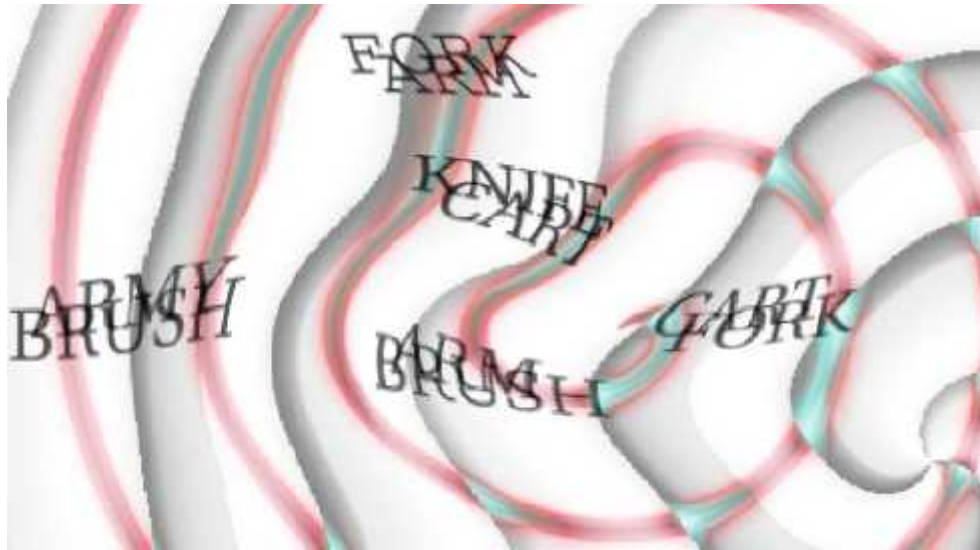
- ❑ Original motivation?
  - Automated bots stuffed ballot box in vote for best CS grad school
- ❑ Free email services — spammers like to use bots to sign up for 1000s of email accounts
  - CAPTCHA employed so only humans get accounts
- ❑ Sites that do not want to be automatically indexed by search engines
  - CAPTCHA would force human intervention

# CAPTCHA: Rules of the Game

- ❑ Easy for most humans to pass
- ❑ Difficult or impossible for machines to pass
  - **Even with access to CAPTCHA software**
- ❑ From Trudy's perspective, the only unknown is a random number
  - Similar to Kerckhoffs' Principle
- ❑ Good to have different CAPTCHAs in case someone cannot pass one type
  - E.g., blind person could not pass visual CAPTCHA

# Do CAPTCHAs Exist?

- ❑ Test: Find 2 words in the following



- ❑ Easy for most humans
- ❑ A (difficult?) OCR problem for computer
  - OCR — Optical Character Recognition



# CAPTCHAs

- ❑ Current types of CAPTCHAs
  - Visual — like previous example
  - Audio — distorted words or music
- ❑ No text-based CAPTCHAs
  - Maybe this is impossible...

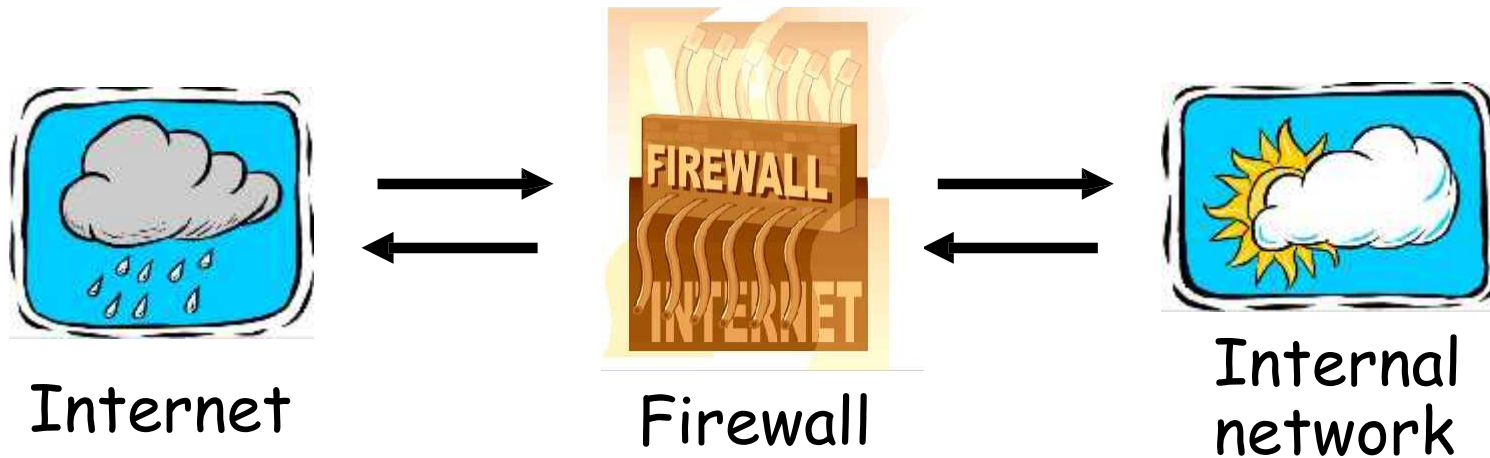
# CAPTCHA's and AI

- ❑ OCR is a challenging AI problem
    - Hardest part is the **segmentation problem**
    - Humans good at solving this problem
  - ❑ Distorted sound makes good CAPTCHA
    - Humans also good at solving this
  - ❑ Hackers who break CAPTCHA have solved a hard AI problem (such as OCR)
    - So, putting hacker's effort to good use!
  - ❑ Other ways to defeat CAPTCHAs???
- See the excellent paper, Telling Humans and Computers Apart: How Lazy Cryptographers do AI:  
[http://www.captcha.net/captcha\\_cacm.pdf](http://www.captcha.net/captcha_cacm.pdf)

# Firewalls



# Firewalls



- ❑ Firewall decides what to let in to internal network and/or what to let out
- ❑ **Access control** for the network

# Firewall as Secretary

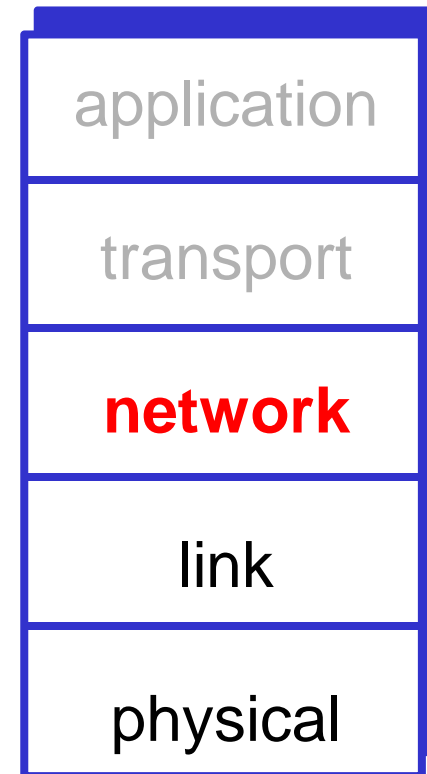
- ❑ A firewall is like a **secretary**
- ❑ To meet with an executive
  - First contact the secretary
  - Secretary decides if meeting is important
  - So, secretary filters out many requests
- ❑ You want to meet chair of CS department?
  - Secretary does some filtering
- ❑ You want to meet the POTUS?
  - Secretary does lots of filtering

# Firewall Terminology

- ❑ No standard firewall terminology
- ❑ Types of firewalls
  - **Packet filter** — works at network layer
  - **Stateful packet filter** — transport layer
  - **Application proxy** — application layer
- ❑ Lots of other terms often used
  - E.g., “deep packet inspection”

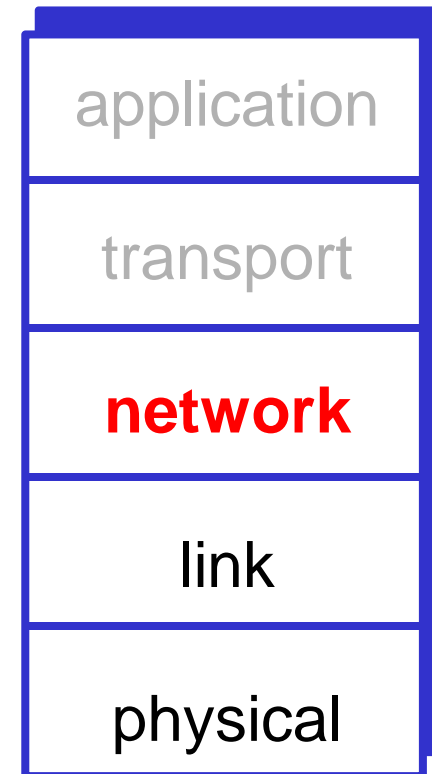
# Packet Filter

- ❑ Operates at network layer
- ❑ Can filters based on...
  - Source IP address
  - Destination IP address
  - Source Port
  - Destination Port
  - Flag bits (SYN, ACK, etc.)
  - Egress or ingress



# Packet Filter

- ❑ Advantages?
  - Speed
- ❑ Disadvantages?
  - No concept of state
  - Cannot see TCP connections
  - Blind to application data





# Packet Filter

- ❑ Configured via Access Control Lists (ACLs)
  - Different meaning than at start of Authorization lecture

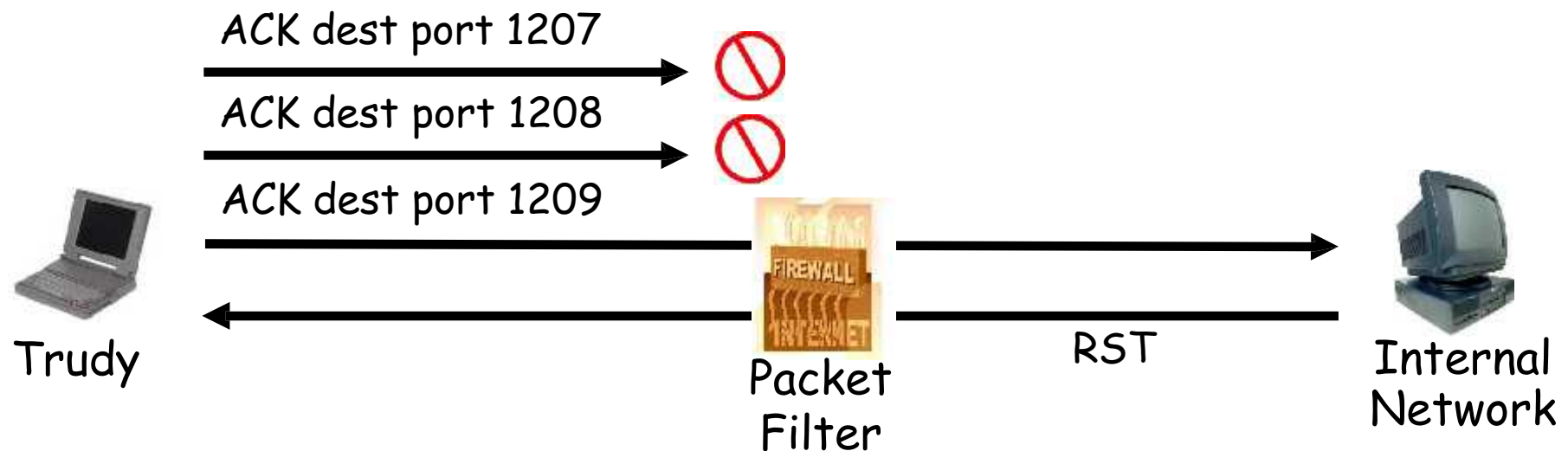
Action	Source IP	Dest IP	Source Port	Dest Port	Protocol	Flag Bits
Allow	Inside	Outside	Any	80	HTTP	Any
Allow	Outside	Inside	80	> 1023	HTTP	ACK
Deny	All	All	All	All	All	All

- ❑ Q: Intention?
- ❑ A: Restrict traffic to Web browsing

# TCP ACK Scan

- ❑ Attacker scans for open ports thru firewall
  - Port scanning is *first step* in many attacks
- ❑ Attacker sends packet with ACK bit set, **without** prior 3-way handshake
  - Violates TCP/IP protocol
  - ACK packet pass thru packet filter firewall
  - Appears to be part of an ongoing connection
  - RST sent by recipient of such packet

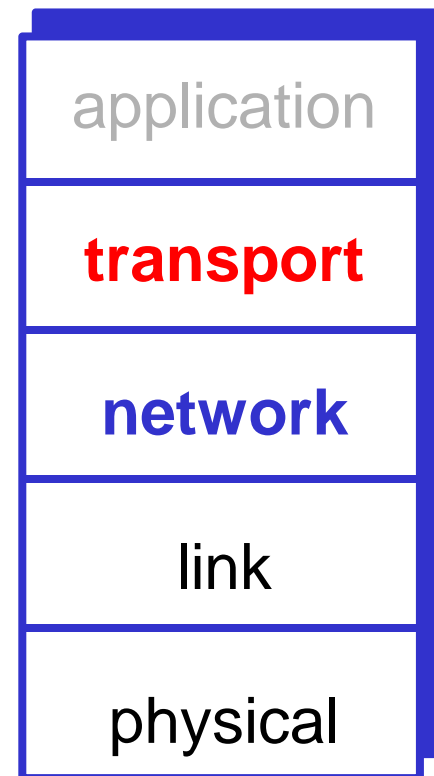
# TCP ACK Scan



- ❑ Attacker knows port 1209 open thru firewall
- ❑ A **stateful packet filter** can prevent this
  - Since scans not part of established connections

# Stateful Packet Filter

- ❑ Adds **state** to packet filter
- ❑ Operates at transport layer
- ❑ ***Remembers*** TCP connections, flag bits, etc.
- ❑ Can even remember UDP packets (e.g., DNS requests)



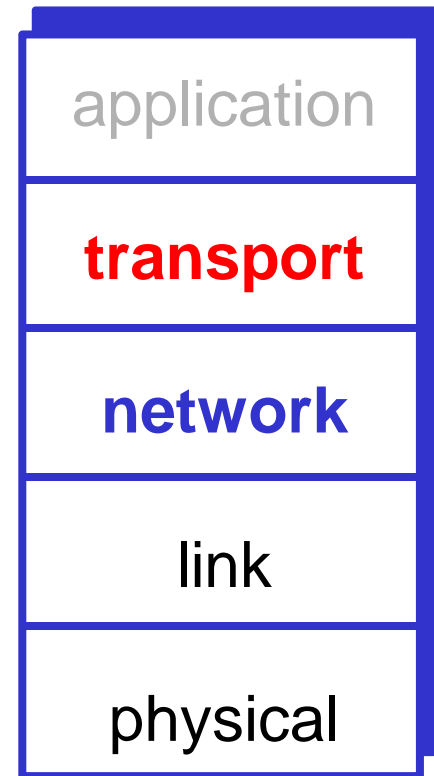
# Stateful Packet Filter

## ❑ Advantages?

- Can do everything a packet filter can do plus...
- Keep track of ongoing connections (e.g., prevents TCP ACK scan)

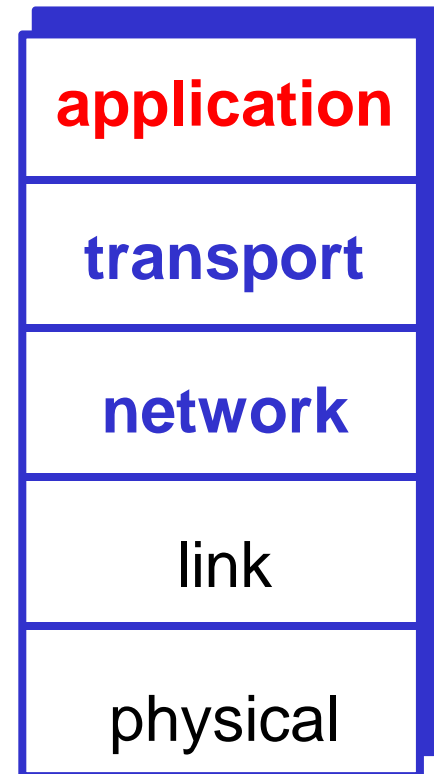
## ❑ Disadvantages?

- Cannot see application data
- Slower than packet filtering



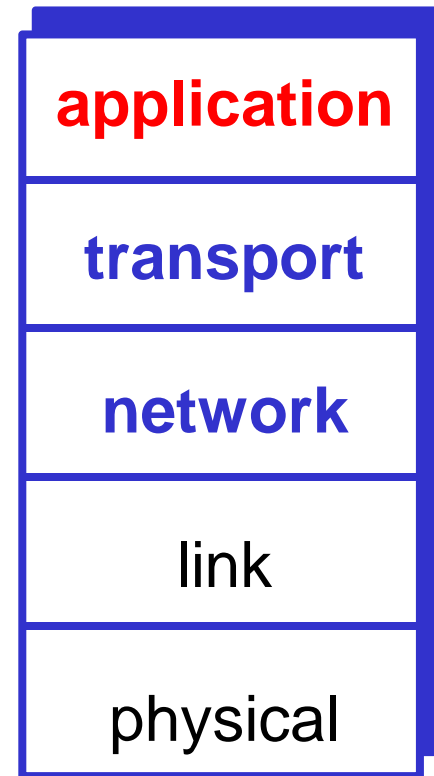
# Application Proxy

- ❑ A **proxy** is something that acts on your behalf
- ❑ Application proxy looks at incoming application data
- ❑ Verifies that data is safe before letting it in



# Application Proxy

- ❑ Advantages?
  - Complete view of connections and applications data
  - Filter bad data at application layer (viruses, Word macros)
- ❑ Disadvantages?
  - Speed



# Application Proxy

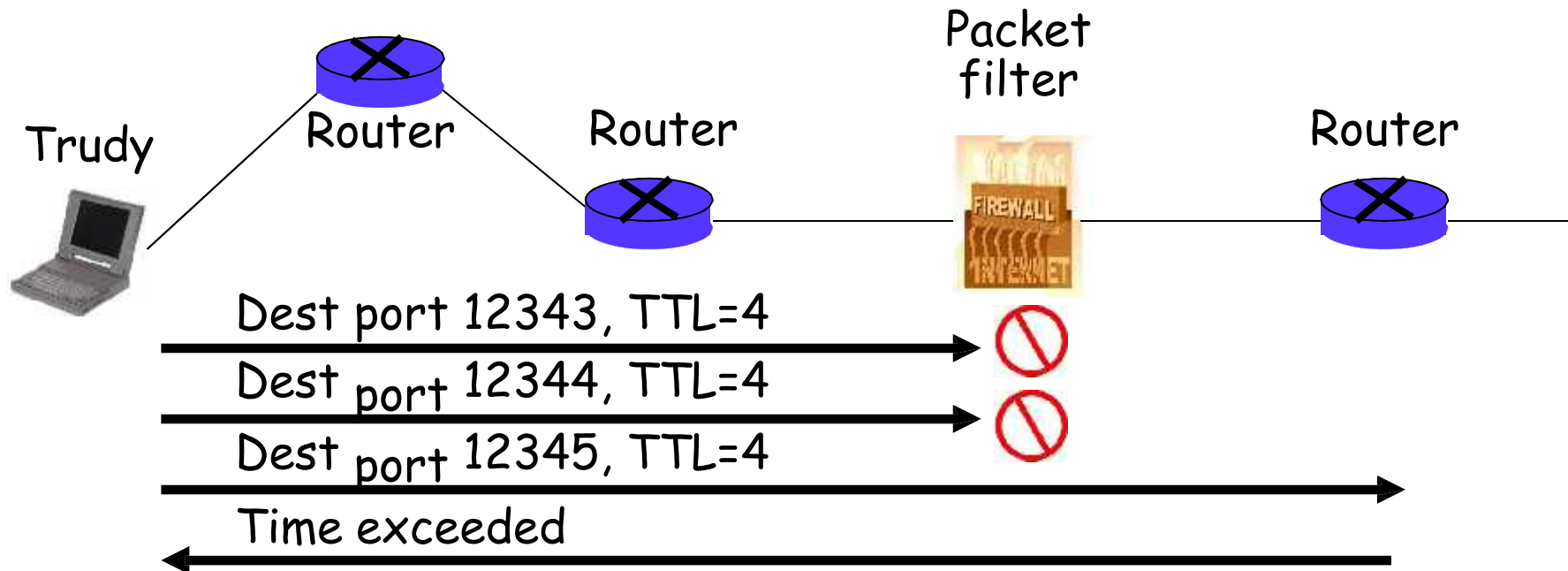
- ❑ Creates a new packet before sending it thru to internal network
- ❑ Attacker must talk to **proxy** and convince it to forward message
- ❑ Proxy has complete view of connection
- ❑ Prevents some scans stateful packet filter cannot — next slides



# Firewalk

- ❑ Tool to scan for open ports thru firewall
- ❑ Attacker knows IP address of firewall and IP address of one system inside firewall
  - Set TTL to 1 more than number of hops to firewall, and set destination port to N
- ❑ If firewall allows data on port N thru firewall, get ***time exceeded*** error message
  - Otherwise, no response

# Firewalk and Proxy Firewall



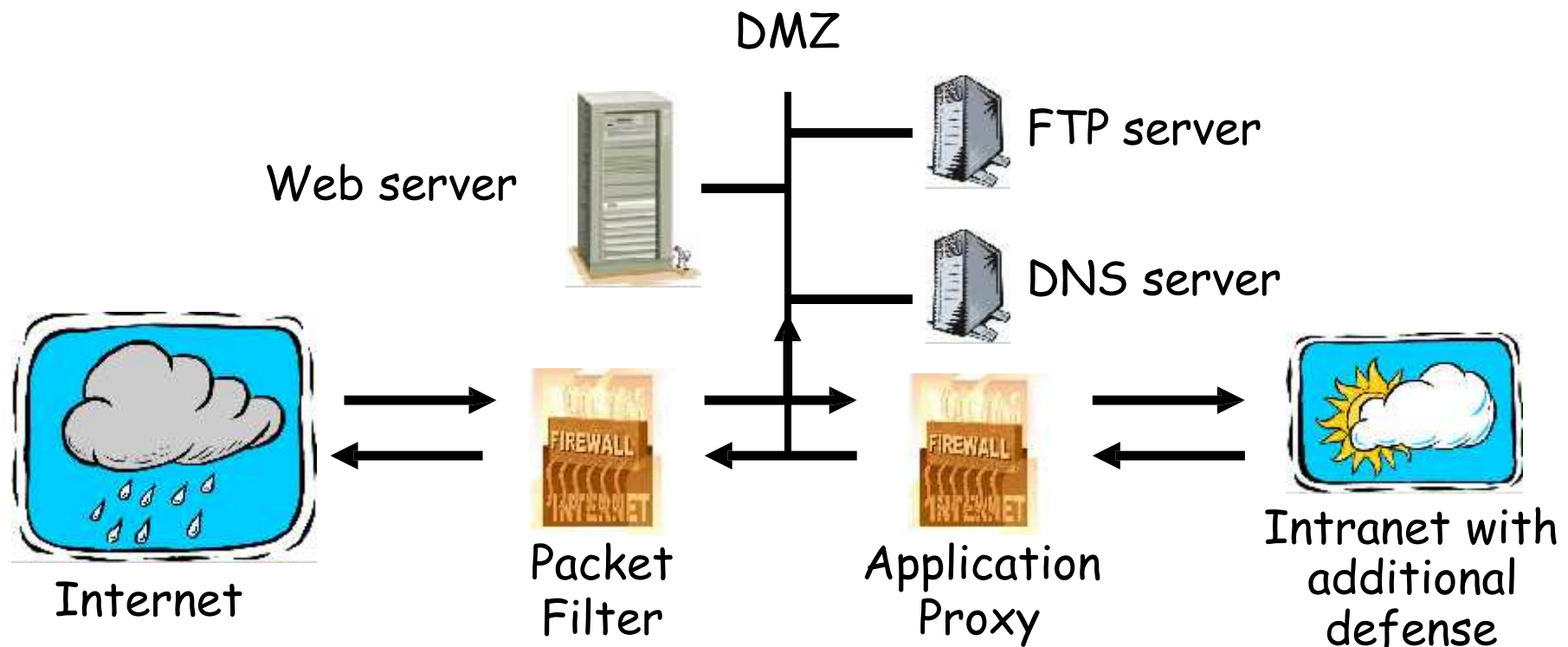
- ❑ This will **not** work thru an application proxy (why?)
- ❑ The proxy creates a new packet, destroys old TTL

# Deep Packet Inspection

- ❑ Many buzzwords used for firewalls
  - One example: **deep packet inspection**
- ❑ What could this mean?
- ❑ Look into packets, but don't really "process" the packets
  - Like an application proxy, but faster

# Firewalls and Defense in Depth

## □ Typical network security architecture



# Intrusion Detection Systems

# Intrusion Prevention

- ❑ Want to keep bad guys out
- ❑ **Intrusion prevention** is a traditional focus of computer security
  - Authentication is to prevent intrusions
  - Firewalls a form of intrusion prevention
  - Virus defenses aimed at intrusion prevention
  - Like locking the door on your car

# Intrusion Detection

- ❑ In spite of intrusion prevention, bad guys will sometime get in
- ❑ Intrusion detection systems (**IDS**)
  - Detect attacks in progress (or soon after)
  - Look for unusual or suspicious activity
- ❑ IDS evolved from log file analysis
- ❑ IDS is currently a **hot** research topic
- ❑ How to respond when intrusion detected?
  - We don't deal with this topic here...

# Intrusion Detection Systems

- ❑ Who is likely intruder?
  - May be outsider who got thru firewall
  - May be evil insider
- ❑ What do intruders do?
  - Launch well-known attacks
  - Launch variations on well-known attacks
  - Launch new/little-known attacks
  - “Borrow” system resources
  - Use compromised system to attack others. etc.



# IDS

- ❑ Intrusion detection **approaches**
  - Signature-based IDS
  - Anomaly-based IDS
- ❑ Intrusion detection **architectures**
  - Host-based IDS
  - Network-based IDS
- ❑ Any IDS can be classified as above
  - In spite of marketing claims to the contrary!

# Host-Based IDS

- ❑ Monitor activities on hosts for
  - Known attacks
  - Suspicious behavior
- ❑ Designed to detect attacks such as
  - Buffer overflow
  - Escalation of privilege, ...
- ❑ Little or no view of network activities

# Network-Based IDS

- ❑ Monitor activity on the network for...
  - Known attacks
  - Suspicious network activity
- ❑ Designed to detect attacks such as
  - Denial of service
  - Network probes
  - Malformed packets, etc.
- ❑ Some overlap with firewall
- ❑ Little or no view of host-base attacks
- ❑ Can have both host and network IDS

# Signature Detection Example

- ❑ Failed login attempts may indicate password cracking attack
- ❑ IDS could use the rule “N failed login attempts in M seconds” as **signature**
- ❑ If N or more failed login attempts in M seconds, IDS warns of attack
- ❑ Note that such a warning is specific
  - Admin knows what attack is suspected
  - Easy to verify attack (or false alarm)

# Signature Detection

- ❑ Suppose IDS warns whenever  $N$  or more failed logins in  $M$  seconds
  - Set  $N$  and  $M$  so false alarms not common
  - Can do this based on “normal” behavior
- ❑ But, if Trudy knows the signature, she can try  $N - 1$  logins every  $M$  seconds...
- ❑ Then signature detection slows down Trudy, but might not stop her

# Signature Detection

- ❑ Many techniques used to make signature detection more robust
- ❑ Goal is to detect “almost” signatures
- ❑ For example, if “about”  $N$  login attempts in “about”  $M$  seconds
  - Warn of possible password cracking attempt
  - What are reasonable values for “about”?
  - Can use statistical analysis, heuristics, etc.
  - Must not increase false alarm rate too much

# Signature Detection

- ❑ Advantages of signature detection:
  - Simple
  - Detect known attacks
  - Know which attack at time of detection
  - Efficient (if there are reasonable number of signatures)
- ❑ Disadvantages of signature detection:
  - Signature files must be kept up to date
  - Number of signatures may become large
  - Can only detect known attacks
  - Variation on known attack may not be detected

# Anomaly Detection

- ❑ Anomaly detection systems look for unusual or abnormal behavior
- ❑ There are (at least) two challenges:
  - What is normal for this system?
  - How “far” from normal is abnormal?
- ❑ No avoiding statistics here!
  - **mean** defines normal
  - **variance** gives distance from normal to abnormal



# How to Measure Normal?

- ❑ How to measure normal?
  - Must measure during “representative” behavior
  - Must not measure during an attack...
  - ...or else attack will seem normal!
  - Normal is statistical **mean**
  - Must also compute **variance** to have any reasonable idea of abnormal

# How to Measure Abnormal?

- ❑ Abnormal is relative to some “normal”
  - Abnormal indicates possible attack
- ❑ Statistical discrimination techniques include
  - Bayesian statistics
  - Linear discriminant analysis (LDA)
  - Quadratic discriminant analysis (QDA)
  - Neural nets, hidden Markov models (HMMs), etc.
- ❑ Fancy modeling techniques also used
  - Artificial intelligence
  - Artificial immune system principles
  - Many, many, many others

# Anomaly Detection (1)

- ❑ Suppose we monitor use of three commands:  
***open, read, close***
- ❑ Under normal use we observe Alice:  
open, read, close, open, open, read, close, ...
- ❑ Of the six possible ordered pairs, we see four pairs are normal for Alice,  
(open,read), (read,close), (close,open), (open,open)
- ❑ Can we use this to identify unusual activity?

# Anomaly Detection (1)

- ❑ We monitor the use of the three commands:  
***open, read, close***
- ❑ If the ratio of abnormal to normal pairs is “too high”, warn of possible attack
- ❑ Could improve this approach by
  - Also use expected frequency of each pair
  - Use more than two consecutive commands
  - Include more commands/behavior in the model
  - More sophisticated statistical discrimination

# Anomaly Detection (2)

- Over time, Alice has accessed file  $F_n$  at rate  $H_n$

$H_0$	$H_1$	$H_2$	$H_3$
.10	.40	.40	.10

- Recently, “Alice” has accessed  $F_n$  at rate  $A_n$

$A_0$	$A_1$	$A_2$	$A_3$
.10	.40	.30	.20

- Is this normal use for Alice?
- We compute  $S = (H_0 - A_0)^2 + (H_1 - A_1)^2 + \dots + (H_3 - A_3)^2 = .02$ 
  - We consider  $S < 0.1$  to be normal, so this is normal
- How to account for use that varies over time?

## Anomaly Detection (2)

- ❑ To allow “normal” to adapt to new use, we update averages:  $H_n = 0.2A_n + 0.8H_n$
- ❑ In this example,  $H_n$  are updated...  
 $H_2 = .2 * .3 + .8 * .4 = .38$  and  $H_3 = .2 * .2 + .8 * .1 = .12$
- ❑ And we now have

$H_0$	$H_1$	$H_2$	$H_3$
.10	.40	.38	.12

# Anomaly Detection (2)

- The updated long term average is

$H_0$	$H_1$	$H_2$	$H_3$
.10	.40	.38	.12

- Suppose new observed rates...

$A_0$	$A_1$	$A_2$	$A_3$
.10	.30	.30	.30

- Is this normal use?
- Compute  $S = (H_0 - A_0)^2 + \dots + (H_3 - A_3)^2 = .0488$ 
  - Since  $S = .0488 < 0.1$  we consider this normal
- And we again update the long term averages:  
$$H_n = 0.2A_n + 0.8H_n$$

# Anomaly Detection (2)

- ❑ The starting averages were:

$H_0$	$H_1$	$H_2$	$H_3$
.10	.40	.40	.10

- ❑ After 2 iterations, averages are:

$H_0$	$H_1$	$H_2$	$H_3$
.10	.38	.364	.156

- ❑ Statistics slowly evolve to match behavior
- ❑ This reduces false alarms for SA
- ❑ But also opens an avenue for attack...
  - Suppose Trudy **always** wants to access  $F_3$
  - Can she convince IDS this is normal for Alice?



## Anomaly Detection (2)

- ❑ To make this approach more robust, must incorporate the variance
- ❑ Can also combine N stats  $S_i$  as, say,  
$$T = (S_1 + S_2 + S_3 + \dots + S_N) / N$$
  
to obtain a more complete view of “normal”
- ❑ Similar (but more sophisticated) approach is used in an IDS known as **NIDES**
- ❑ NIDES combines anomaly & signature IDS

# Anomaly Detection Issues

- ❑ Systems constantly evolve and so must IDS
  - Static system would place huge burden on admin
  - But evolving IDS makes it possible for attacker to (slowly) convince IDS that an attack is normal
  - Attacker may win simply by “going slow”
- ❑ What does “abnormal” really mean?
  - Indicates there may be an attack
  - Might not be any specific info about “attack”
  - How to respond to such vague information?
  - In contrast, signature detection is very specific

# Anomaly Detection

## ❑ Advantages?

- Chance of detecting unknown attacks

## ❑ Disadvantages?

- Cannot use anomaly detection alone...
- ...must be used with signature detection
- Reliability is unclear
- May be subject to attack
- Anomaly detection indicates “something unusual”, but lacks specific info on possible attack

# Anomaly Detection: Conclusion

- ❑ Anomaly-based IDS is active research topic
- ❑ Many security experts have high hopes for its ultimate success
- ❑ Often cited as key future security technology
- ❑ Hackers are not convinced!
  - Title of a talk at Defcon: “Why Anomaly-based IDS is an Attacker’s Best Friend”
- ❑ Anomaly detection is difficult and tricky

# Access Control Summary

- Authentication and authorization
  - Authentication — who goes there?
    - Passwords — something you know
    - Biometrics — something you are (you are your key)
    - Something you have

# Access Control Summary

- ❑ Authorization — are you allowed to do that?
  - Access control matrix/ACLs/Capabilities
  - MLS/Multilateral security
  - BLP/Biba
  - Covert channel
  - Inference control
  - CAPTCHA
  - Firewalls
  - IDS

# Coming Attractions...

- ❑ Security protocols
  - Generic authentication protocols
  - SSH
  - SSL
  - IPSec
  - Kerberos
  - WEP
  - GSM
- ❑ We'll see lots of crypto applications in the protocol