COMP 3355
Application Layer Security

K.P. Chow

University of Hong Kong

AUTHENTICATION

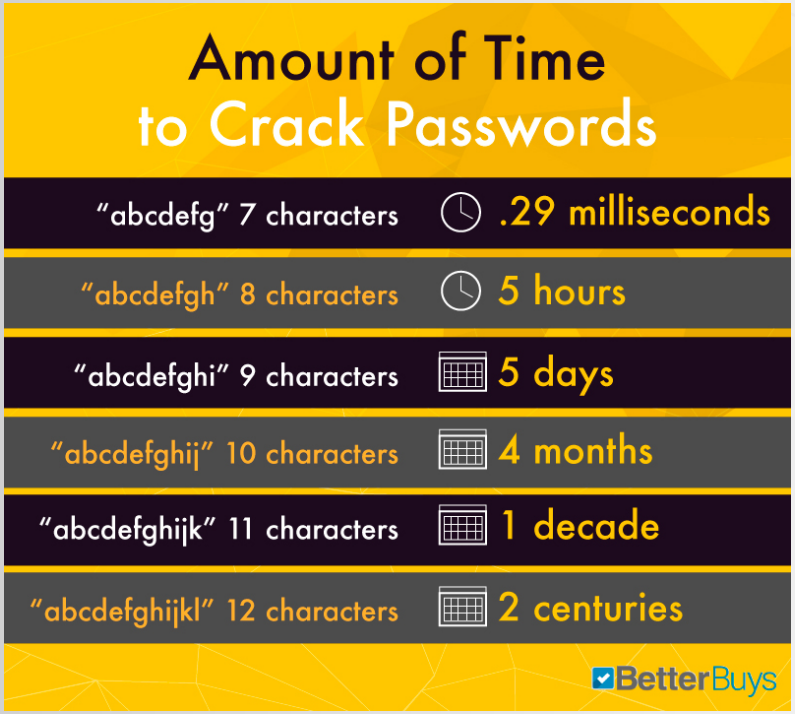
Identification and Authentication

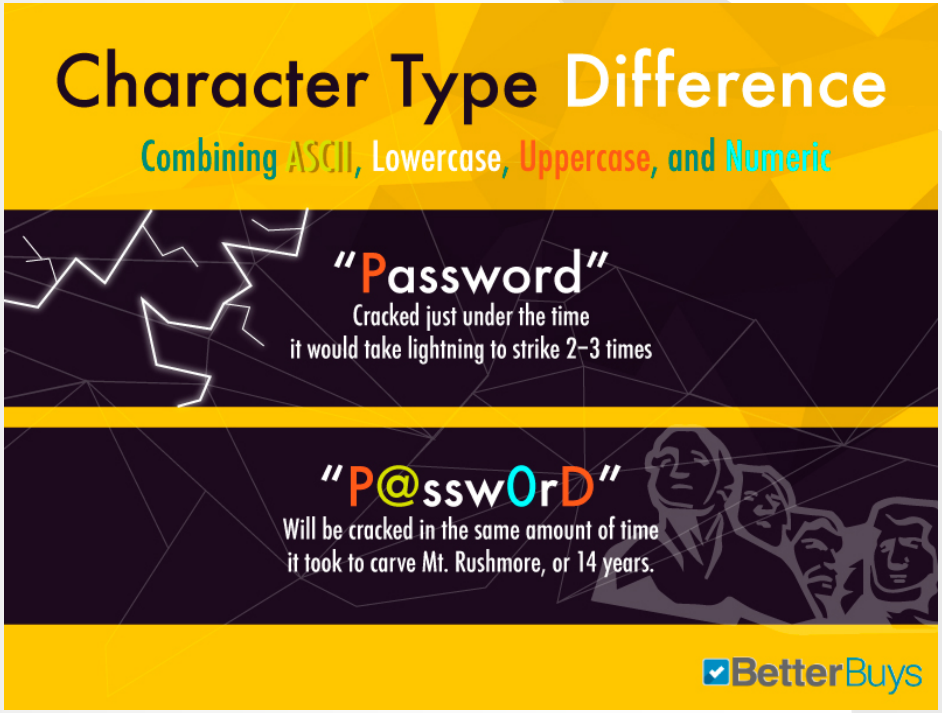
- Username and Password
  - Most common scheme to login to a computer system, carried out in 2 steps:
  1. Identification: enter the user name
  2. Authentication: enter the password
- Other authentication techniques:
  1. Something you know, e.g. password
  2. Something you have, e.g. smartcard
  3. Something you are, e.g. biometric
- Definitions:
  - Subject: a user
  - Credential: the required proof needed by the system to validate the identity of the user, e.g. the password
  - Principal: a name associated with a subject, e.g. username
  - Subject can have multiple names, e.g. different username on different machine

Password Study

Group	Password Criteria	Try to crack the passwords
A	Passwords consists of >= 6 characters, with at least 1 non-letter	30% easy to crack (passwords easy to remember)
B	Passwords based on passphrases	10% were cracked (passwords easy to remember)
C	Passwords consists of 8 randomly selected characters	10% were cracked (passwords difficult to remember)

Year 2015


Amount of Time
to Crack Passwords
"abcdefg" 7 characters
.29 milliseconds
"abcdefgh" 8 characters
5 hours
"abcdefghi" 9 characters
5 days
"abcdefghij" 10 characters
4 months
"abcdefghijk" 11 characters
1 decade
"abcdefghijkl" 12 characters
2 centuries
BetterBuys


Character Type Difference
Combining ASCII, Lowercase, Uppercase, and Numeric
"Password"
Cracked just under the time
it would take lightning to strike 2–3 times
"P@ssw0rD"
Will be cracked in the same amount of time
it took to carve Mt. Rushmore, or 14 years.
BetterBuys

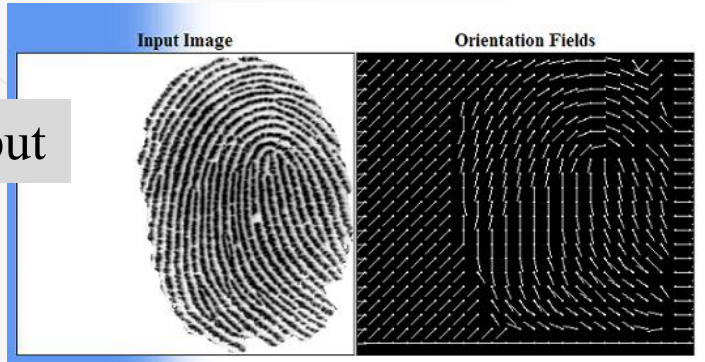
Biometric – Fingerprint

- Enrollment/Registration : recording of fingerprint info
  - Fingerprint is ‘captured’ by a device
  - ‘features’ are extracted (based on some mathematical models)
  - ‘features’ are recorded in a database
- Verification / Identification: checking of a fingerprint
  - Fingerprint is ‘captured’, ‘features’ are extracted
  - ‘features’ are compared with those in database
    - Known-subject matching : match with one stored fingerprint features record
    - Unknown-subject matching: match all records to find records with similar features (more on investigation, less popular)

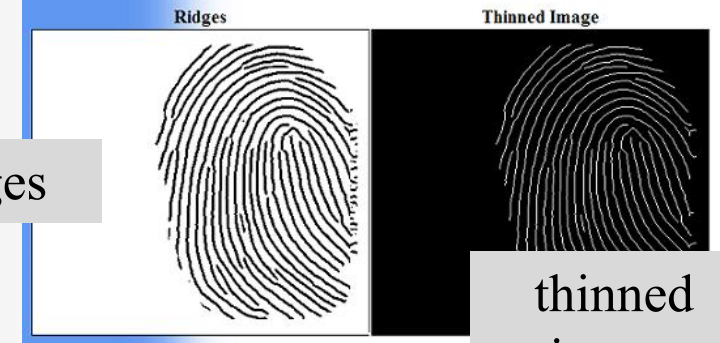
What is a feature?

Fingerprint Scanner

input



ridges

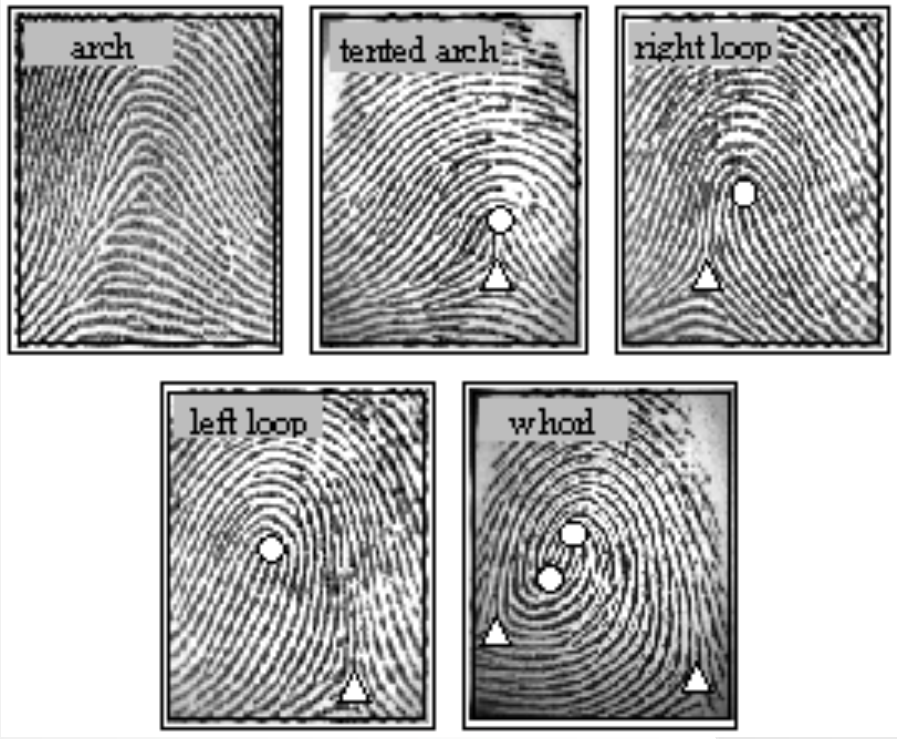


thinned image

minutiae features

Fingerprint Global Features

- Global features: patterns
  - aggregate characteristics of ridges
- 3 basic patterns: arch, loop, whorl
  - Arch: ridges enter from one side of the finger, rise in the center forming an arc, and then exit the other side
  - Loop: ridges enter from one side of a finger, form a curve, and tend to exit from the same side they enter
  - Whorl: ridges form circularly around a central point on the finger



Are we using these in automatic fingerprint recognition?

Fingerprint Local Features

- Local features: minutiae
- Major Minutia features of fingerprint ridges are: ridge ending, bifurcation, and short ridge (or dot).

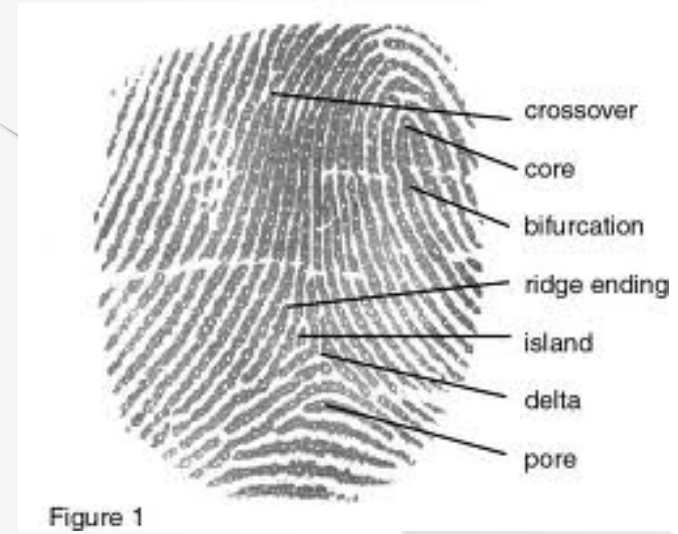
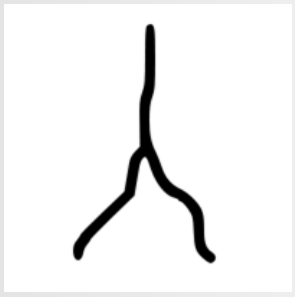


Figure 1

ridge ending (termination)

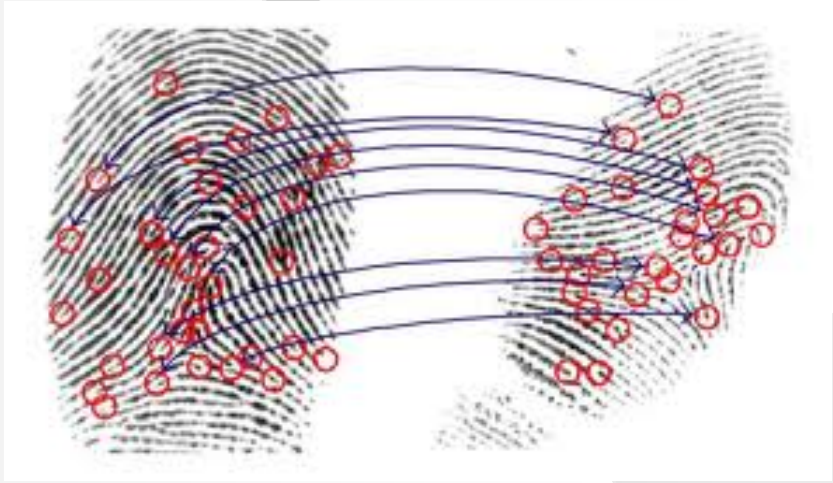
bifurcation



short ridge

Fingerprint Minutiae Matching

- Minutiae are extracted from the 2 fingerprints and stored as sets of points in the 2-dimensional plane
- Each minutia is stored as a triplet {x,y, θ}, where (x,y) is the minutia is the location coordinate and θ is the minutia angle
- Formulated as a "point pattern matching" problem and processed by pattern recognition algorithm



template

input fingerprint

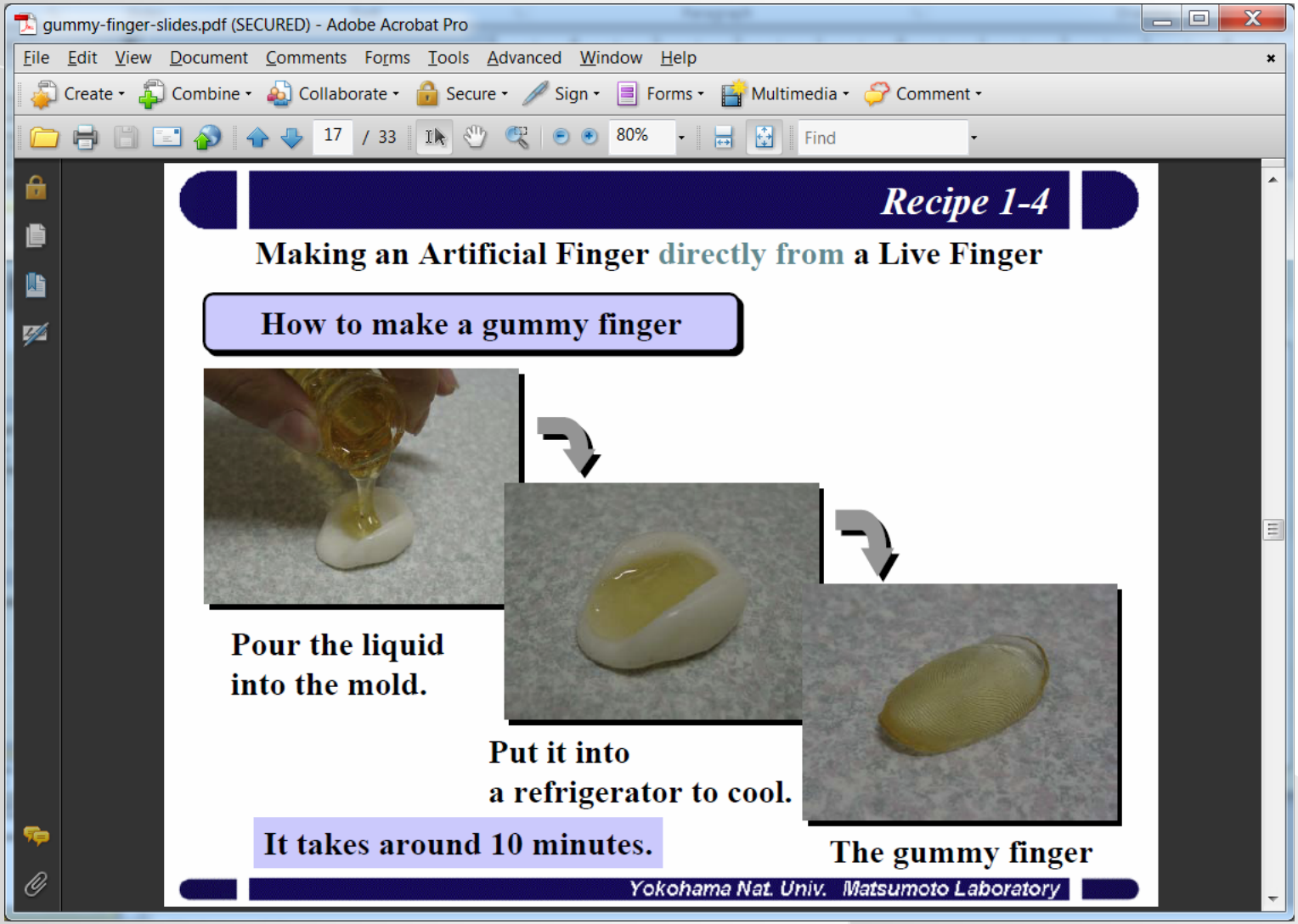
The Guidance Note

HONG KONG MONETARY AUTHORITY
香港金融管理局

- Should implement proper techniques to authenticate the identity and authority of their customers and their e-banking systems, and to ensure the integrity of information transmitted over networks
- Customer authentication can be achieved by any one of the following methods: (i) something a customer knows (e.g., passwords); (ii) something a customer has (e.g., digital certifi
token); (iii) something)

Stronger customer authentication combines at least two of the above methods

Artificial Finger



HOST-BASED INTRUSION DETECTION

Intrusion Detection – Motivations & Basis

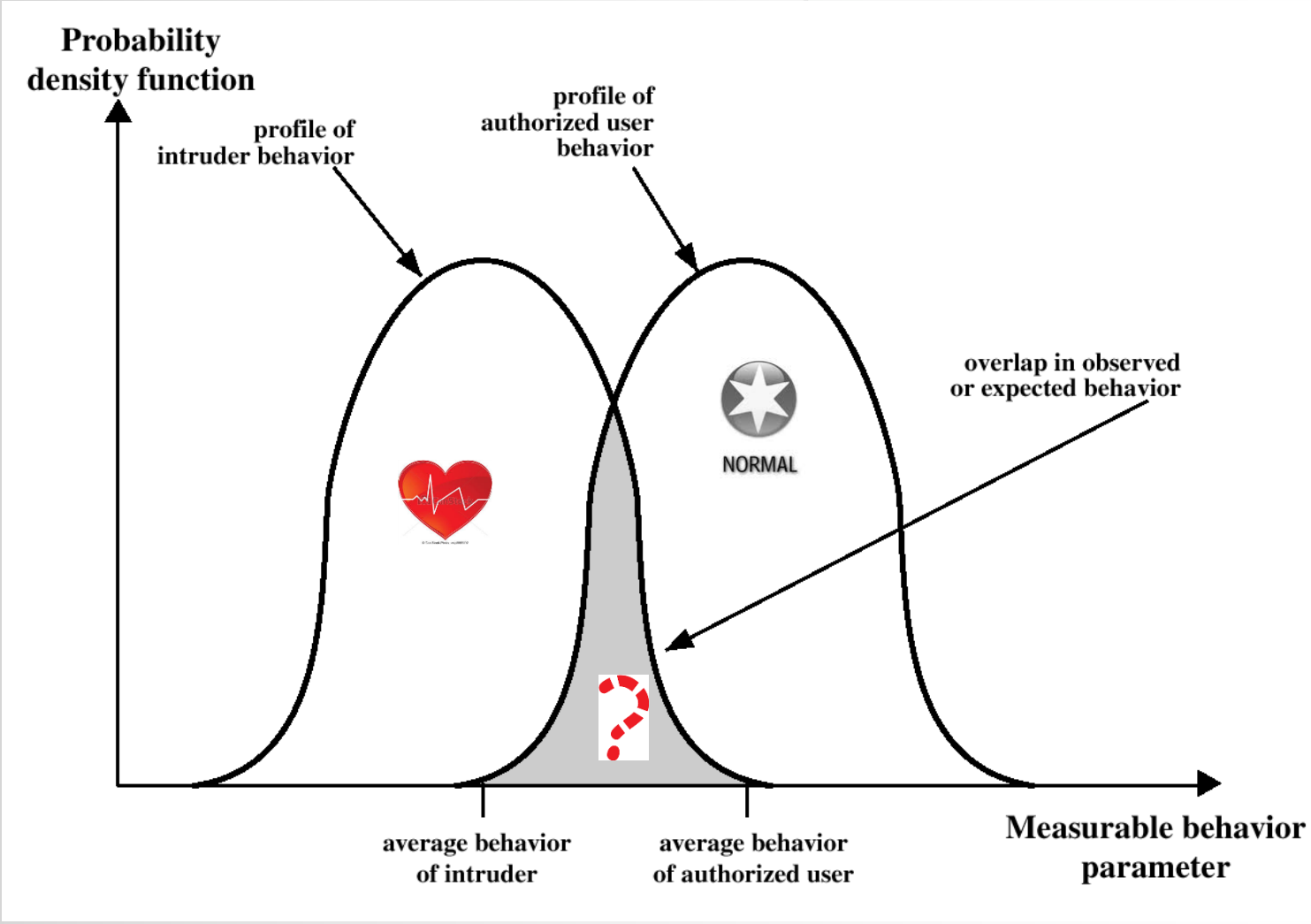
Motivations:...

- Collection of information about intrusion techniques that can be used to strengthen the intrusion prevention facility.

Intrusion detection basis:

"The behavior of intruders differs from that of a legitimate user in a quantifiable way."

Profiles of Behavior of Intruders and Authorized Users



Intruders and Intrusions

- Intruders:
  - Masquerader: an unauthorized principal who penetrates the OS's access controls to gain a legitimate user's permission
  - Misfeasor: a legitimate user who either access resources for which he is not authorized, or misuses access to resources that he is authorized to access
  - Clandestine user: an individual who seizes supervisory control of the OS and uses it to evade auditing
- Intrusions:
  - Attempts to copy password file regularly
  - Suspicious RPC requests periodically
  - Multiple attempts to connect to nonexistent "bait" machines

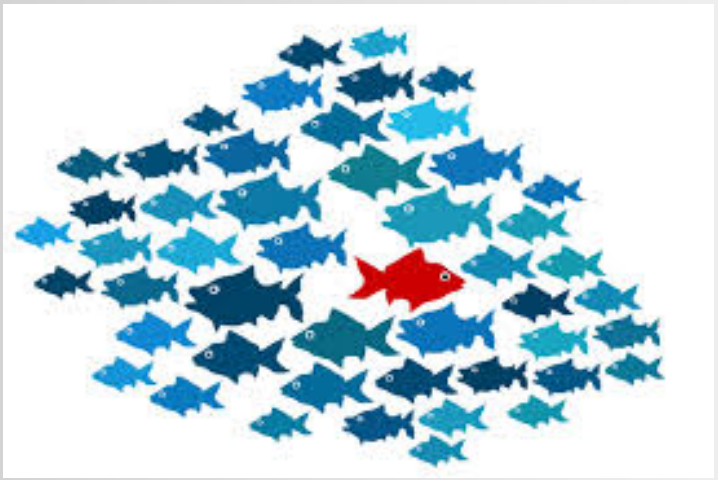
Host-Based Intrusion Detection – Audit Records

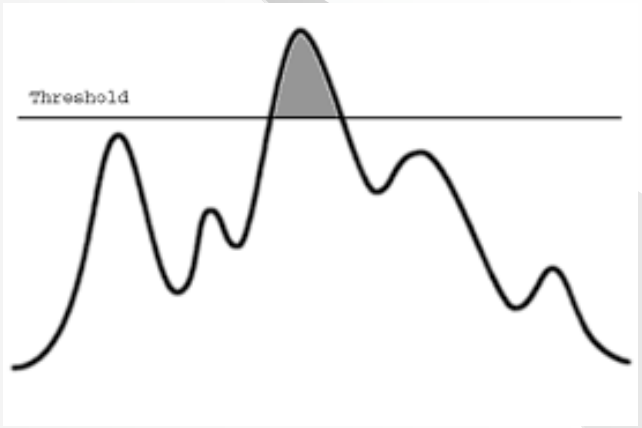
- Traditional approach to ID: protecting the OS on the basis of audit records, such as audit trails, log files, …
- Native audit record is to collect information on user activity, mainly for accounting purpose
- Detection specific audit record is an entry generated when a security-critical operation is performed by a subject on an object
- Typical audit record:

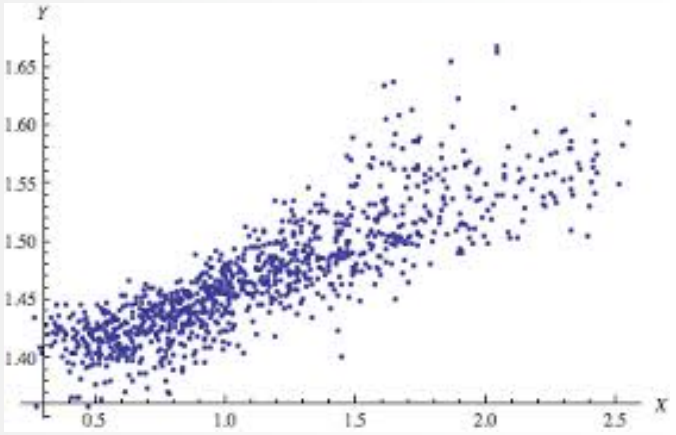
Subject	Action	Object	Exception Condition	Resource Usage	Time Stamp
KPC	read	secret.txt	read violation	RECORDS=0	20110310: 23:12
Hui	execute	Copy.exe	write violation	RECORDS=0	20110311: 02:55

Statistical Intrusion Detection

- Commonly used in host-based IDS
- 3 methods: anomaly detection, threshold detection, correlation methods







Statistical Methods for Intrusion Detection

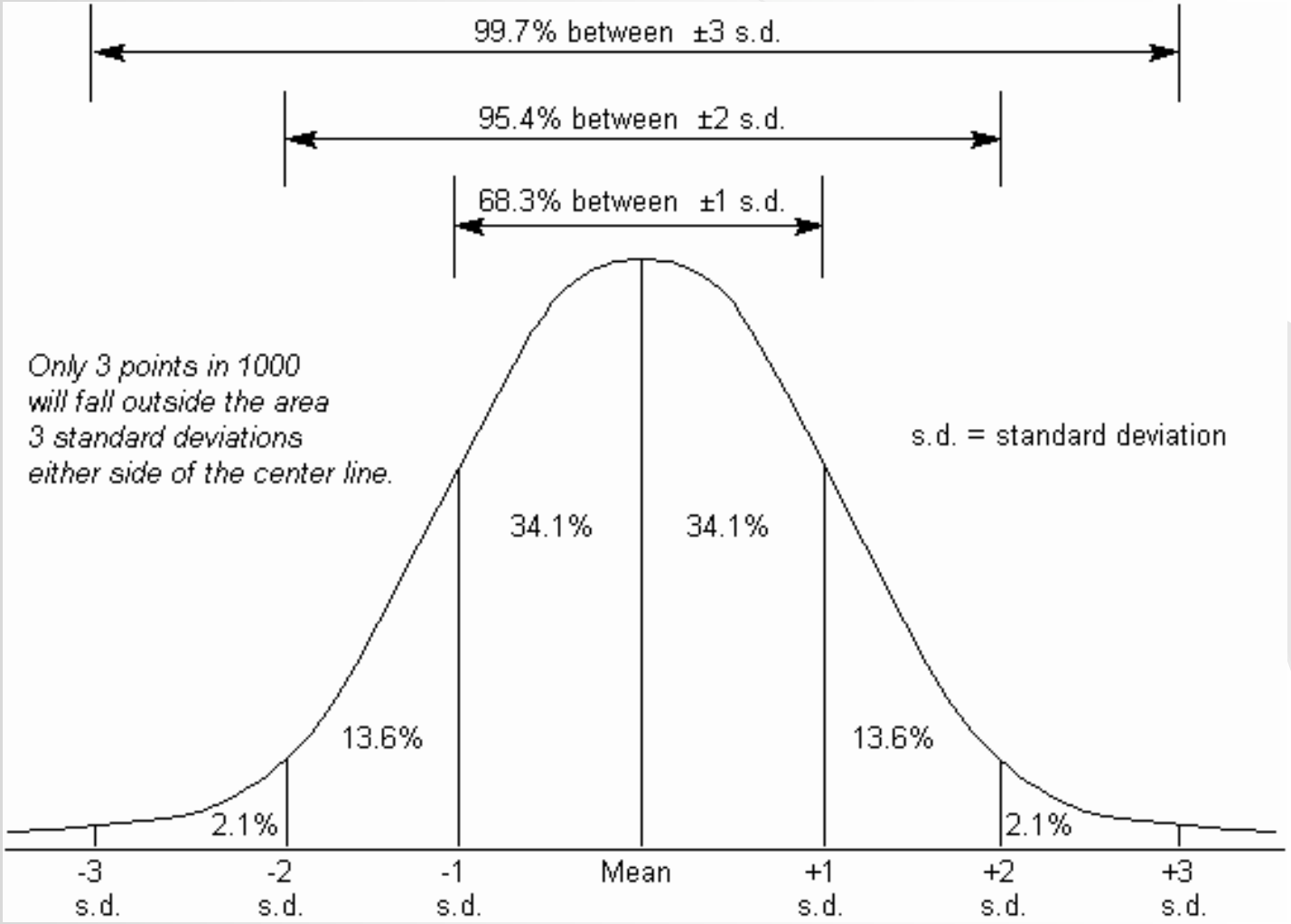
Uses statistical methods to generate values and patterns that are typical for a particular system:

- Counter: a count of certain event types over a period of time, e.g. number of password failure
- Gauge: measure the current value of some entities, e.g. number of logical connections assigned to a user application, number of messages generated by a process
- Interval timer: length of time between 2 events, e.g. time between successive logins to an account
- Resource utilization: quantity of resources consumed in a specified period, e.g. number of pages printed by a user session

Statistical Measurements

- Mean and standard deviation: average and variability
- Multivariate analysis: measure the correlation
- Markov process: measure the transition among states, e.g. transition between certain commands
- Time series: focus on time intervals, looking for sequence of events that happen too rapidly or too slowly
- Operational: based on a judgment of what is considered abnormal, instead of based on past records, e.g. large number of login attempts in a short period of time

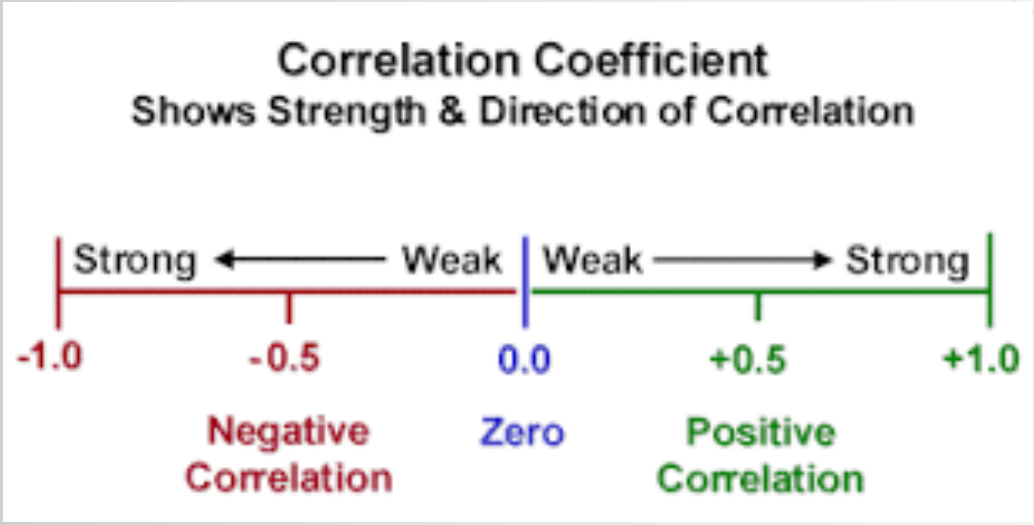
Mean and Standard Deviation



Mean and Standard Deviation Model

- Known events: $x_1, \ldots x_n$
- Mean and standard deviation (SD):
  - sum = $x_1 + \ldots + x_n$
  - *sumsquares* = $x_1^2 + \ldots + x_n^2$
  - *mean = sum/ n*
  - *stdev = sqrt (sumsquares / (n+1) – mean^2)*
- New observation x_{n+1} is defined to be abnormal if it falls outside a *confidence interval* that is *d* standard deviations from the mean for some parameter *d*:
  - *mean + d * stdev*
- By Chebyshev's inequality, the probability of a value falling outside this interval is at most $1 / d^2$
  - E.g., for d = 4, the probability is at most 0.0625

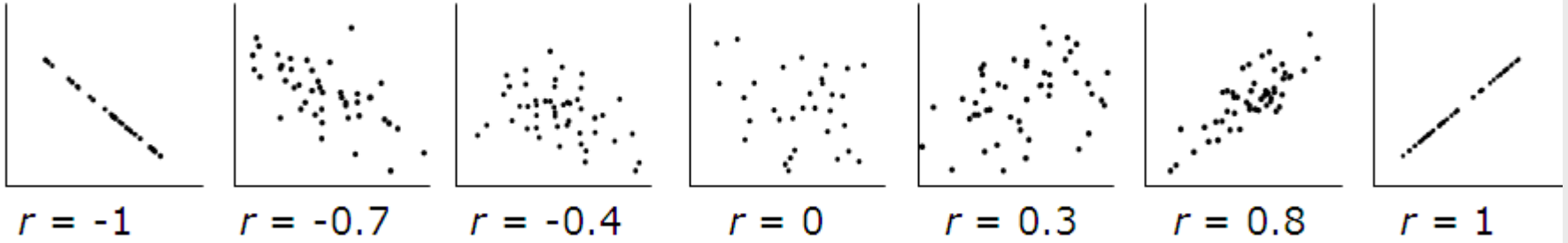
Correlation Coefficient



$$r = \frac{\sum_i (x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_i (x_i - \overline{x})^2} \sqrt{\sum_i (y_i - \overline{y})^2}}$$

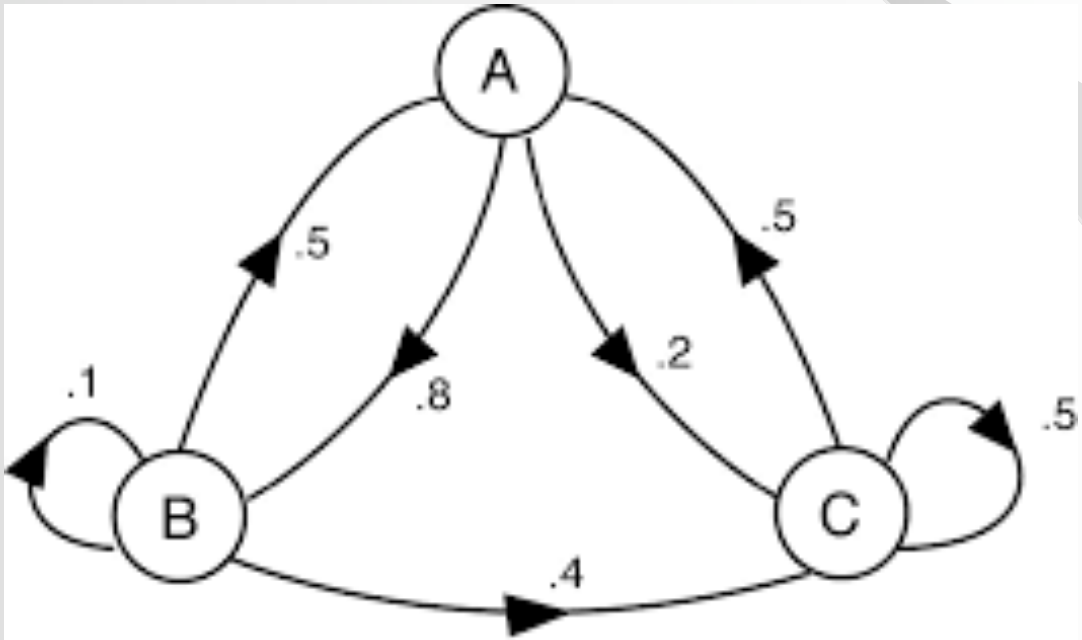
Get the correlation coefficient (r) from your calculator or computer

- r has a value between -1 and +1:



Markov Process

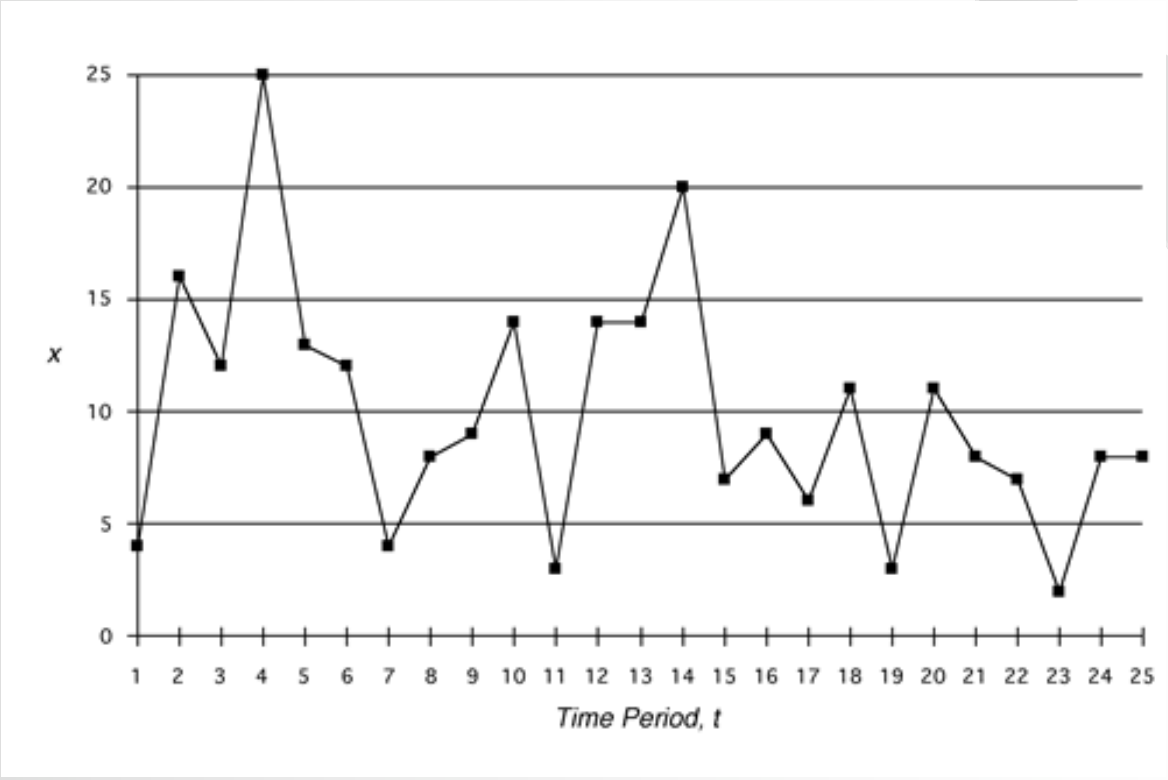
- Measure the transition among states, e.g. transition between certain commands



Markov graph of transiton probabilites between states A, B and C

Time Series

- Focus on time intervals, looking for sequence of events that happen too rapidly or too slowly



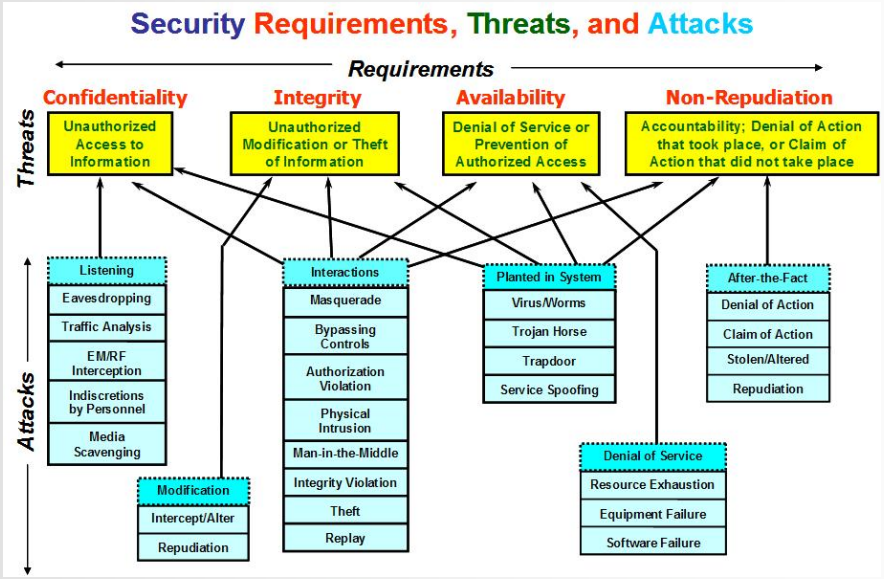
Comparison of Host and Network based IDS

	Network IDS	Host IDS
Advantage	■ No need to install any tools to application or server ■ Can detect attack attempts to number of servers ■ OS independent	■ More accurate observation of attack ■ Can be specially designed for application
Disadvantage	■ More false alarm results ■ Generically design for network environment	■ Need to install agents to communicate with server ■ Collect attack attempts from individual devices ■ OS dependent

SECURITY TESTING

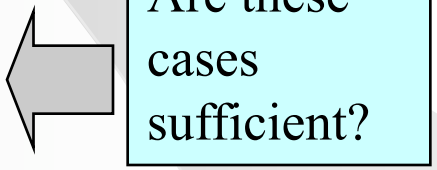
Security Testing

- Traditional software testing addresses the application requirements and verifies that the use case scenarios have been implemented
  - It does not test the scenarios and actions that an application is not supposed to allow
  - Attack use cases and scenarios based on attack patterns are not considered
- What are appropriate "security requirements" for a system?



Simple Security Requirement

- Security requirement:
  - Authentication: the application should use password to authenticate the user
- Test cases:
  - Login successful: user login with a valid password
  - Login failure: user fails to login with an invalid passwords
  - Multiple login failure:
  - Password change:



“Real” Security Testing

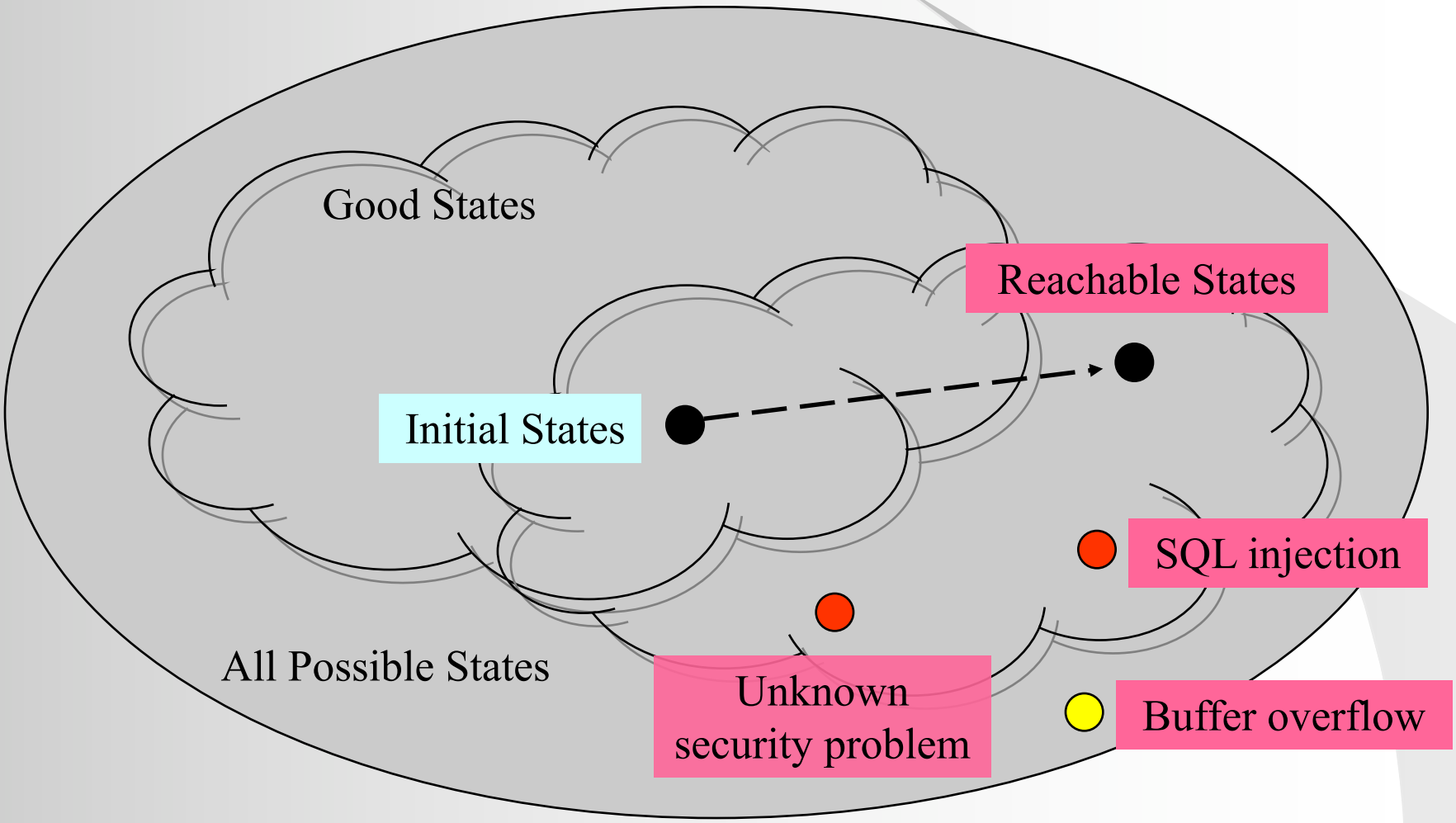
- Security testing: checking that some features appears to fail, e.g.
  - The tester is able to spoof another user’s identity
  - The tester can tamper the data
  - The tester can view data he should not have access to
  - The tester can deny service to other users
  - …

Security Testing Example

- Security testing is to prove that defensive mechanisms work correctly rather than proving that the functions work properly, e.g. for the bank account example, one attack pattern is called “SQL injection”:
  - An attacker input SQL commands as part of normal user input and the SQL commands become part of a SQL query to a SQL server
  - A simple way is see if user input is being used to build a dynamically generated SQL query is to input a single-quote character as part of the input
  - If the output looks like originating from a database server, it is likely that the system is not protecting against SQL injection attacks, e.g.

    Microsoft OLE DB provider for ODBC Drivers error ‘80040e07’

Security Testing – State Spaces View



Security Testing – Attack Testing

- Attack testing forces a program to perform actions on invalid or malicious data to reveal what the program could allow an attacker to do
- Attack pattern is a series of repeatable steps that can be applied to simulate an attack against the security of a system
  - To ensure potential vulnerabilities are considered
  - To highlight areas which need to be considered for security testing

Security Testing – Attack Testing Example

- Examples
  - SQL injection: verify that user input from an attacker is not allowed to manipulate the back-end database
  - Cross-site scripting: verify an attack that can cause an attacker's script to execute in a victim's Web browser is not possible
  - Unexpected input: verify that errors are handled correctly so that the program safely recovers from unexpected input

Security Testing Process (Howard and LeBlanc)

1. Identify the component interfaces (using data flow analysis and threat model)
2. Rank the interfaces by potential vulnerability
3. Construct test cases according to well known security problems (STRIDE):
   - Spoofing, Tampering, Repudiation, Information disclosure, Denial of service, Elevation of privilege

STRIDE (1)

- Spoofing identity (S):
  - Client spoofing: allows an attacker to pose as another user, e.g. insecure basic HTTP Authentication
  - Server spoofing: allows a rogue server to pose as a valid server, e.g. DNS spoofing
- Tampering with data (T): malicious modification of data, e.g.
  - Change data in a file with weak ACL (Everyone)
  - Unauthorized update of data in the database
- Repudiation (R): user denies performing an action while the other party cannot prove the existence of the action, e.g.
  - A customer purchased a book from an online bookshop and later denied such purchase, and the bookshop was unable to prove that the purchase was done by the customer

STRIDE (2)

- Information disclosure (I): the exposure of information to individuals who are not supposed to have access to it, e.g.
  - Read a file in a shared file system without the corresponding access right
  - Read data in transit between 2 computers, such as the username/password using basic HTTP Authentication
- Denial of service (D): denies service to valid users, e.g.
  - A web server temporarily unavailable or unusable
- Elevation of privilege (E): an unprivileged user gains privileged access and thereby has sufficient access to compromise or destroy an entire system, e.g.
  - Trojan runs with “root” privileges and open a backdoor for the hacker

Penetration Test