Modelling and Intrusion Detection

Joseph Spring

Discussion – Modelling and Intrusion Detection

1. What is the Problem?
2. Feature List
3. Reduced Feature List
4. Set up the Model
5. Solve the Model
6. Compare with Reality

Discussion – Modelling and Intrusion Detection

What is the Problem?

- To detect the presence of Intruders on a network, … , Distributed system

Feature List

- Features that affect the problem
  - Types of intruder
  - Techniques used by different intruder
  - Approach to be taken for solving the problem

Reduced Feature List

- Smaller number of Types of intruder, techniques, one approach

Discussion – Modelling and Intrusion Detection

Set up the Model

Approach to be taken, use of audit records for historic data, suitability of format in audit records, metrics used, significance level for transgressions …

Solve the Model

- Statistical Anomaly Detection
- Rule Based Intrusion Detection, or …

Compare with Reality

- How effective is the approach with the type of attacker, is the model acceptable?
- The Base-Rate Fallacy
- Distributed Intrusion Detection bottlenecks, data format

Feature List – Intruders

- Significant problems for network security come in the form of hostile/unwanted trespass either by users or software
  - Unauthorised login
  - Unauthorised elevation of privilege
  - Virus, worm, Trojan horse, …
- Two of the most publicised threats are
  - Malware
  - Intruders

Feature List – Intruders

Three Classes of Intruder identified

- Masquerader (generally an outsider)
  - Unauthorised to use computer
  - Penetrates system access controls to exploit a legitimate user's account
- Misfeasor (generally an insider)
  - Legitimate user that accesses data, programs, resources etc; for which either s/he is unauthorised to access or alternatively to misuse that that s/he is authorised to access
- Clandestine User (insider or outsider)
  - Seizes supervisory control of system
  - Uses control to avoid auditing and access controls or to suppress audit collection

Feature List – Intruders

- Benign Intruders
  - People exploring the internet just to see what’s there
- Malevolent Intruders
  - Attempt to read confidential data
  - Unauthorised modification (integrity)
  - Disrupt systems
- Generally two levels of hacker
  - Sophisticated users with very good knowledge of technology (Sophisticated Knowledge)
  - Low level users that use cracking programs with little understanding of how they work (Willing to spend hours probing for weaknesses)

Feature List – Examples of Intrusion

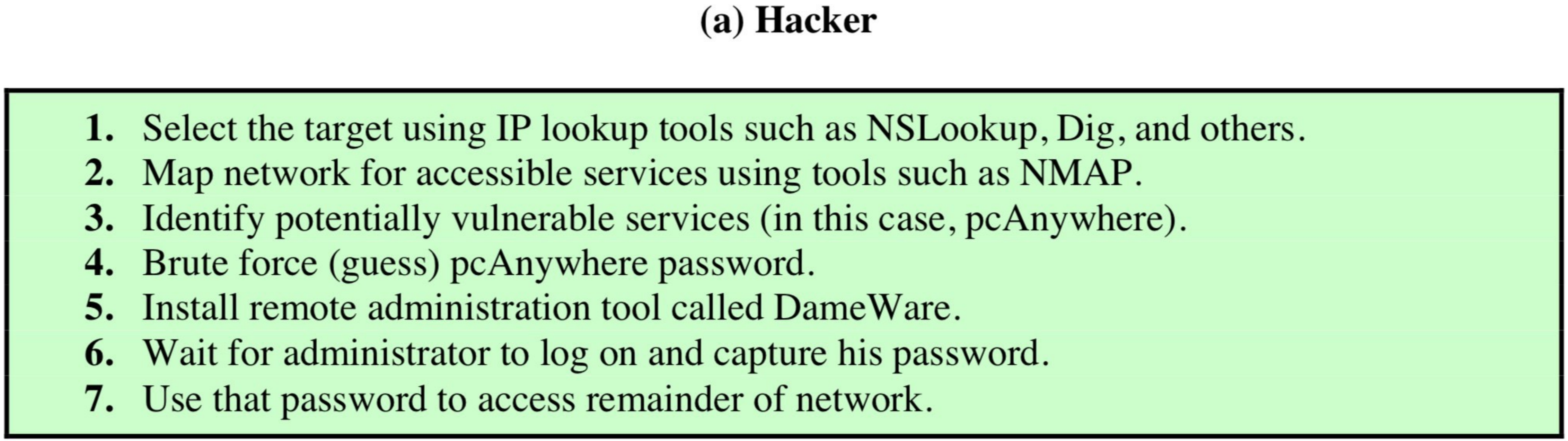
[GRAN04] lists the following examples of intrusion:

- Performing a remote root compromise of an e-mail server
- Defacing a Web server
- Guessing and cracking passwords
- Copying a database containing credit card numbers
- Viewing sensitive data, including payroll records and medical information, without authorization
- Running a packet sniffer on a workstation to capture usernames and passwords

Feature List – Examples of Intrusion

- Using a permission error on an anonymous FTP server to distribute pirated software and music files
- Dialing into an unsecured modem and gaining internal network access
- Posing as an executive, calling the help desk, resetting the executive’s e-mail password, and learning the new password
- Using an unattended, logged-in workstation without permission

Feature List - Intruder Patterns of Behaviour

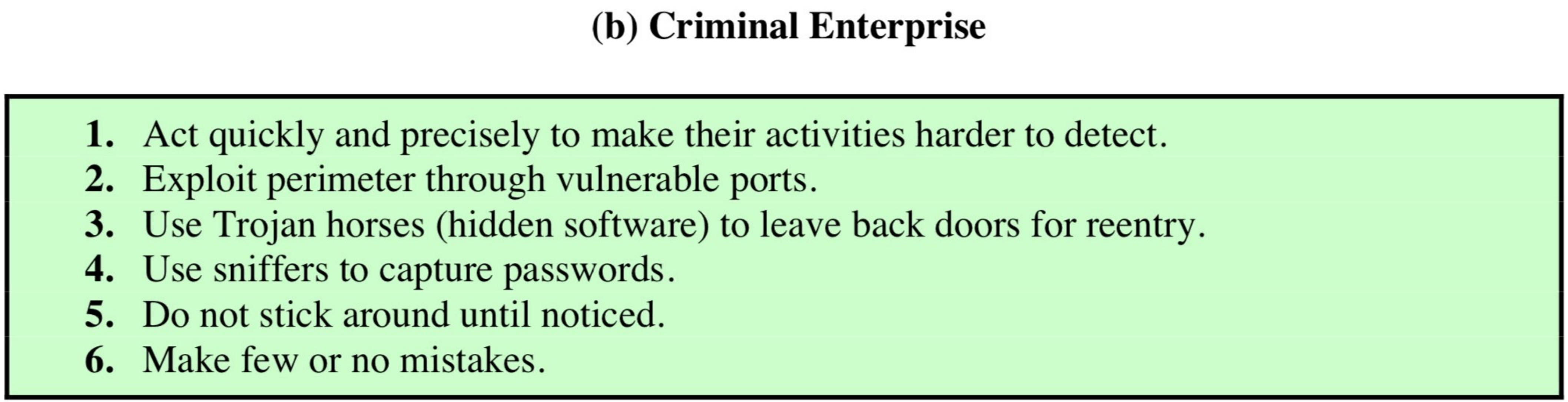


(a) Hacker

1. Select the target using IP lookup tools such as NSLookup, Dig, and others.
2. Map network for accessible services using tools such as NMAP.
3. Identify potentially vulnerable services (in this case, pcAnywhere).
4. Brute force (guess) pcAnywhere password.
5. Install remote administration tool called DameWare.
6. Wait for administrator to log on and capture his password.
7. Use that password to access remainder of network.

1. IP addresses - Indicates location in World
2. pc Anywhere - Remote Desktop connection
3. Dameware (Solar Winds) - remotely access end user computers,.... for remote administration and support.

Feature List - Intruder Patterns of Behaviour



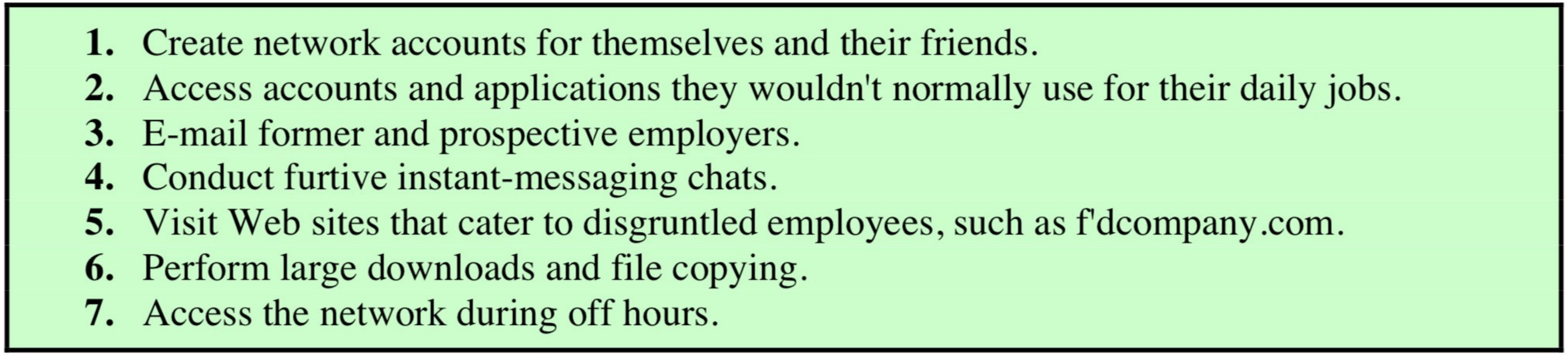
(b) Criminal Enterprise

1. Act quickly and precisely to make their activities harder to detect.
2. Exploit perimeter through vulnerable ports.
3. Use Trojan horses (hidden software) to leave back doors for reentry.
4. Use sniffers to capture passwords.
5. Do not stick around until noticed.
6. Make few or no mistakes.

4- packet sniffer - packet analyser - protocol analyser - network analyser
used to monitor network traffic
eg wireshark.

Feature List - Intruder Patterns of Behaviour

(c) Internal Threat



1. Create network accounts for themselves and their friends.
2. Access accounts and applications they wouldn't normally use for their daily jobs.
3. E-mail former and prospective employers.
4. Conduct furtive instant-messaging chats.
5. Visit Web sites that cater to disgruntled employees, such as f'dcompany.com.
6. Perform large downloads and file copying.
7. Access the network during off hours.

Feature List - Insider Attack: Other Approaches

- Enforce least privilege, only allowing access to the resources employees need to do their job.
- Set logs to see what users access and what commands they are entering.
- Protect sensitive resources with strong authentication. — Multi factor
- Upon termination, delete employee's computer and network access.
- Upon termination, make a mirror image of employee's hard drive before reissuing it. That evidence might be needed if your company information turns up at a competitor.

Feature List - Passwords

1. How big should a password be?
2. What mixture of characters should you use?
3. How are they stored?
4. How do Hackers try to get your passwords?

Feature List - Intrusion Techniques that have been used

1. Try default passwords
2. Exhaustive attempt of all short passwords
3. Try
   - Words in system’s online dictionary
   - List of likely passwords
   - Examples to be found on hackers bulletin boards
4. Collect information on users
   - Full names, name of spouse, children, pictures in office, books in office
   - related to hobbies, …
5. Try user’s phone numbers, social security numbers, room numbers, …
6. Try all legitimate licence plate numbers
7. Use a Trojan horse to bypass restrictions on access
8. Tap the line between a remote user and the host system

Feature List - Intrusion Techniques

- Examples 1 - 6 can be blocked when connecting to host
  - E.g. by rejecting login after say three attempts
  - Intruder has to reconnect to host and try again
  - Impractical to attempt more than a few passwords
- However if an intruder can gain access with say, low level privileges to an encrypted password file then
  - Capture file
  - Carry out analysis (elsewhere) until suitable passwords obtained – i.e. passwords giving access with higher privileges
  - Guessing attacks are considered feasible and highly effective provided
    - They can be attempted automatically
    - Each guess can be verified
    - The guessing process is undetected

Feature List - Intrusion Techniques

- Trojan Horse
  - Difficult to counter
- Example
  - Low privilege user A develops game
  - Invites system operator B to try game in spare time
  - Game contains code to copy B's password (which turns out to be unencrypted but access protected) into A's user account
  - Game running under operators high privilege mode allowed access into password file

Feature List - Intrusion Techniques

- Line Tapping
  - Physical Security
  - Counter via link encryption techniques
- Intrusion Prevention is a systems *first line of defence*
- The *second line of defence* is intrusion detection the focus of much recent research
  - Interest motivated by various consideration:
    - If intrusion detection efficient intruder
      - can be identified and ejected quickly from system minimising damage and facilitating recovery
    - An effective intrusion detection system can serve as an effective deterrent, helping to prevent intrusion
  - Intrusion detection aids collection of information regarding intrusion techniques
    - Helps in development of intrusion detection techniques

Set up Model – Intrusion and Detection

- Intrusion Detection Techniques
  - Based on assumption
    - Intruders behaviour is quantifiably different to a legitimate users behaviour
    - Profiles may be constructed for intruder and user
    - Some overlap between an intruders and users expected but there will be some difference
  - As a consequence
    - a loose interpretation of an intruders behaviour which will catch more intruders may also lead to authorised users being identified as intruders
    - Whilst an attempt to limit the number of incorrect identifications may lead to intruders being mistaken as authorised users
    - An element of compromise and art is therefore to be found in intrusion detection

Qn: Does this lead to our inability to develop intrusion dection s/w with a 100% accuracy when looking for intruders.

Set up Model – Intrusion and Detection

- Masquerader
  - Techniques felt to be particularly useful
    - Construct profile of legitimate user behaviour by observing past history
    - Significant deviation from such behaviour detectable
- Misfeasor
  - More difficult to detect with profile alone
  - Distinction between normal and abnormal behaviour may be small
  - intelligent definition regarding class of conditions that suggest unauthorised use
- Clandestine User
  - Felt to be beyond the scope of automated techniques
- Observations still felt to be applicable today!

Set up Model – Intrusion and Detection

We consider the following techniques:

1. Statistical Anomaly Detection
   - Data collection carried out for legitimate user over period of time
   - Statistical tests applied to observed behaviour to determine, with high level of confidence, if behaviour is that of unauthorised user

   a) Threshold Detection:
   - Involves defining thresholds, independent of user, for the frequency of occurrences of various events

   b) Profile Based:
   - A profile of the activity of each user is developed and used to detect changes in the behaviour of individual accounts

2. Rule-Based Detection
   - Involves attempt to define a set of rules that can be used to decide whether a users behaviour is that of an intruders

   a) Anomaly detection:
   - Rules are developed to detect deviation from previous usage patterns

   b) Penetration Identification:
   - An expert systems approach that searches for suspicious behaviour

Set up Model – Audit Records

- These are fundamental to Intrusion Detection
  - A record of ongoing activity for users is used as input to intrusion detection system
- Two plans:
  - Native Audit Records
    - Nearly all multiuser OS's include accounting software collecting information on user behaviour
    - Advantage; using this data means that additional data collection software is not required
    - Disadvantage; native audit records may not contain required information in suitable form, or at all!
  - Detection Specific Audit Records
    - Collection facility can be implemented to generate audit records as required by intrusion detection scheme
    - Advantage; can be made vendor independent and portable for other systems
    - Disadvantage; extra overhead in parallel running of audit packages

Solve the Problem

Statistical/Stochastic Based Intrusion Detection

Set up Model – Statistical Anomaly Detection

- Threshold Detection
  - Involves counting number of events of a specific event type over an interval of time
  - If count exceeds ‘*reasonable*’ amount then intrusion is assumed
  - By itself, a crude and ineffective detector of even moderately sophisticated attacks
  - Both threshold and time interval must be determined
  - Variability across users leads to false positives and false negatives
  - Simple threshold detectors felt useful if used with other more sophisticated techniques

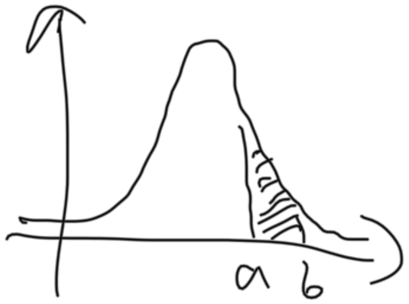
Set up Model – Statistical Anomaly Detection

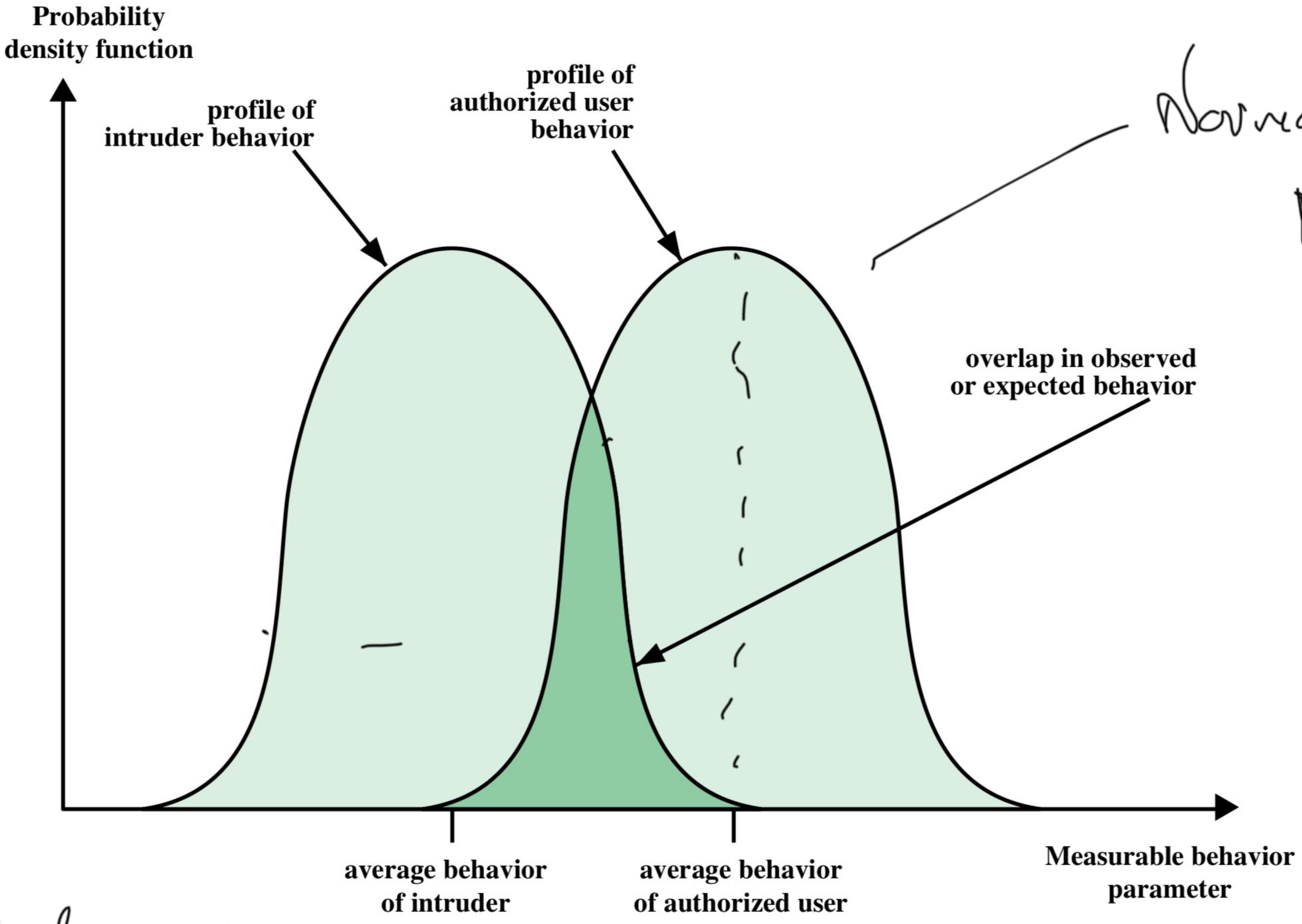
- Profile-Based Anomaly Detection
  - Characterises past behaviour of individual users or related groups of users (establishing expectation)
  - Uses characterisation to detect significant deviations
  - Profile may consist of a collection of parameters
    - Deviation in one parameter may be insufficient to determine intrusion

Set up Model – False Positives

Area under the curves = 1 they represent probabilities.

$P(a<X<b)$





Probability density function
profile of intruder behavior
profile of authorized user behavior
overlap in observed or expected behavior
average behavior of intruder
average behavior of authorized user
Measurable behavior parameter

Normal Distributions

$N(\bar{x}, \sigma^2)$

σ = standard deviation

Set up Model – Statistical Anomaly Detection

- Foundation of approach is the Analysis of Audit Records

  Audit records: Provide input to the intrusion detection function in two ways

  1. Designer must decide on number of quantitative metrics that can be used to measure user behaviour.
     - An analysis of audit records over a period of time can be used to determine activity profile of average user
     - Audit records used to define typical behaviour

  1. Current audit records are the input used to detect intrusion
     - The intrusion detection model analyses new audit records to determine deviation from average behaviour

Set up Model – Statistical and Rule Based Detection

- Statistical Techniques
  - Attempt to define *normal* or *expected* behaviour
  - Are felt to be effective against masqueraders
    - Tend not to mimic behaviour associated with captured accounts
  - Not considered effective against misfeasors
- Rule based approaches
  - attempt to define *proper* behaviour
  - Considered more effective against misfeasors
- A combination of the two approaches is felt to be required in order to be effective against a broad range of attacks

Set up Model – Metrics

Useful Metrics for profile-based intrusion detection:

- Counter
    - Natural number that can increase but not decrease until reset by management action
    - Typically count of particular event types kept over a particular time interval.
    - For example:
        - Number of logins by user in one hour period
        - number of time command executed during user session
        - number of password failures per minute
- Gauge
    - Natural number that can increase or decrease
    - Used to measure current value of entity. For example:
        - Number of logical connections assigned to a user application
        - Number of outgoing messages queued for a user process

Set up Model – Metrics

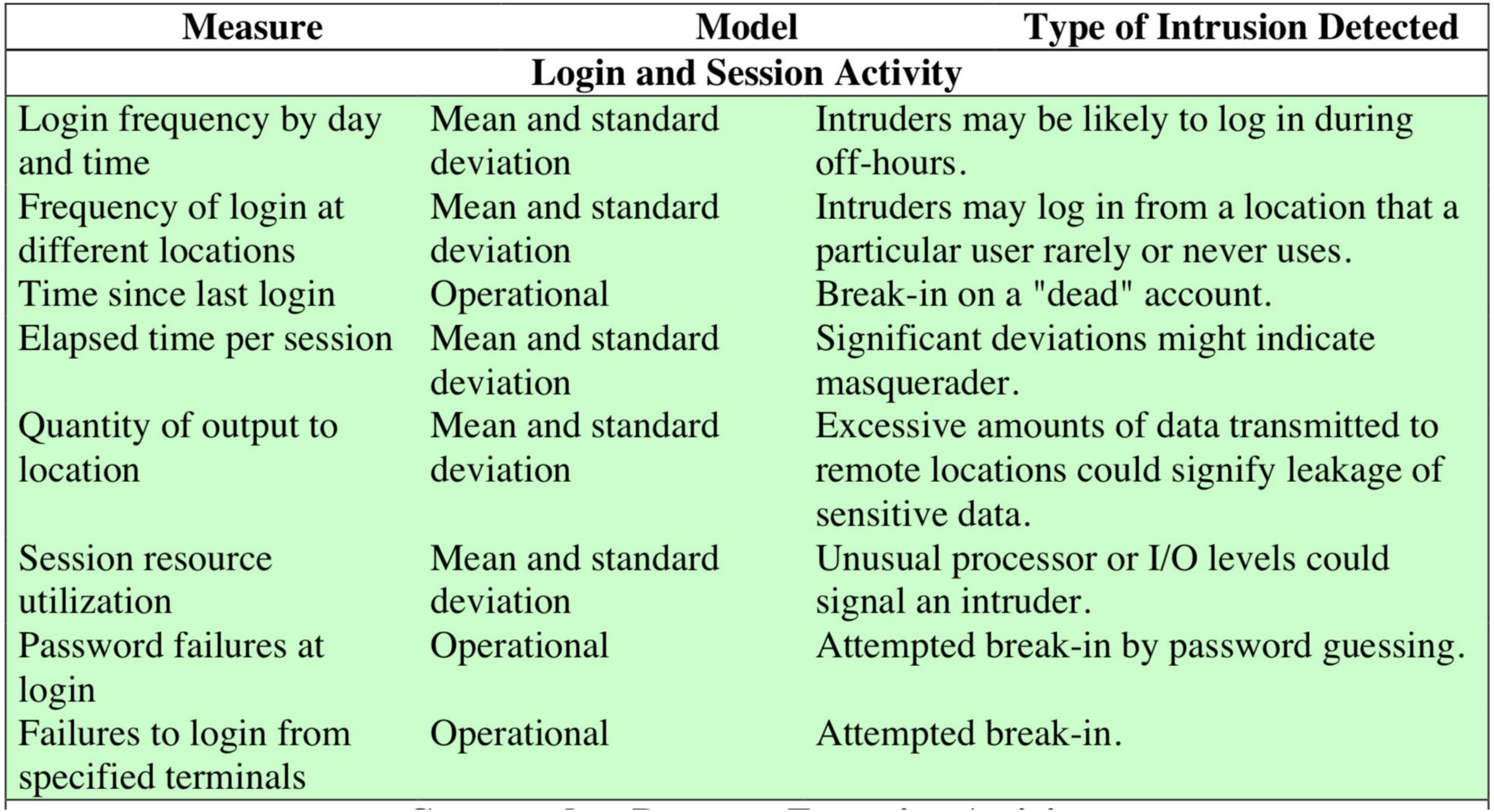
Useful Metrics for profile-based intrusion detection:

- Interval Timer
    - Length of time between two related events. For example:
        - Length of time between successive logins to an account
    - Resource Utilisation
        - Quantity of resources consumed during a specified period. For example:
            - Number of pages printed during a user session
            - Total time consumed b program execution

Solve Model – Metrics and Tests

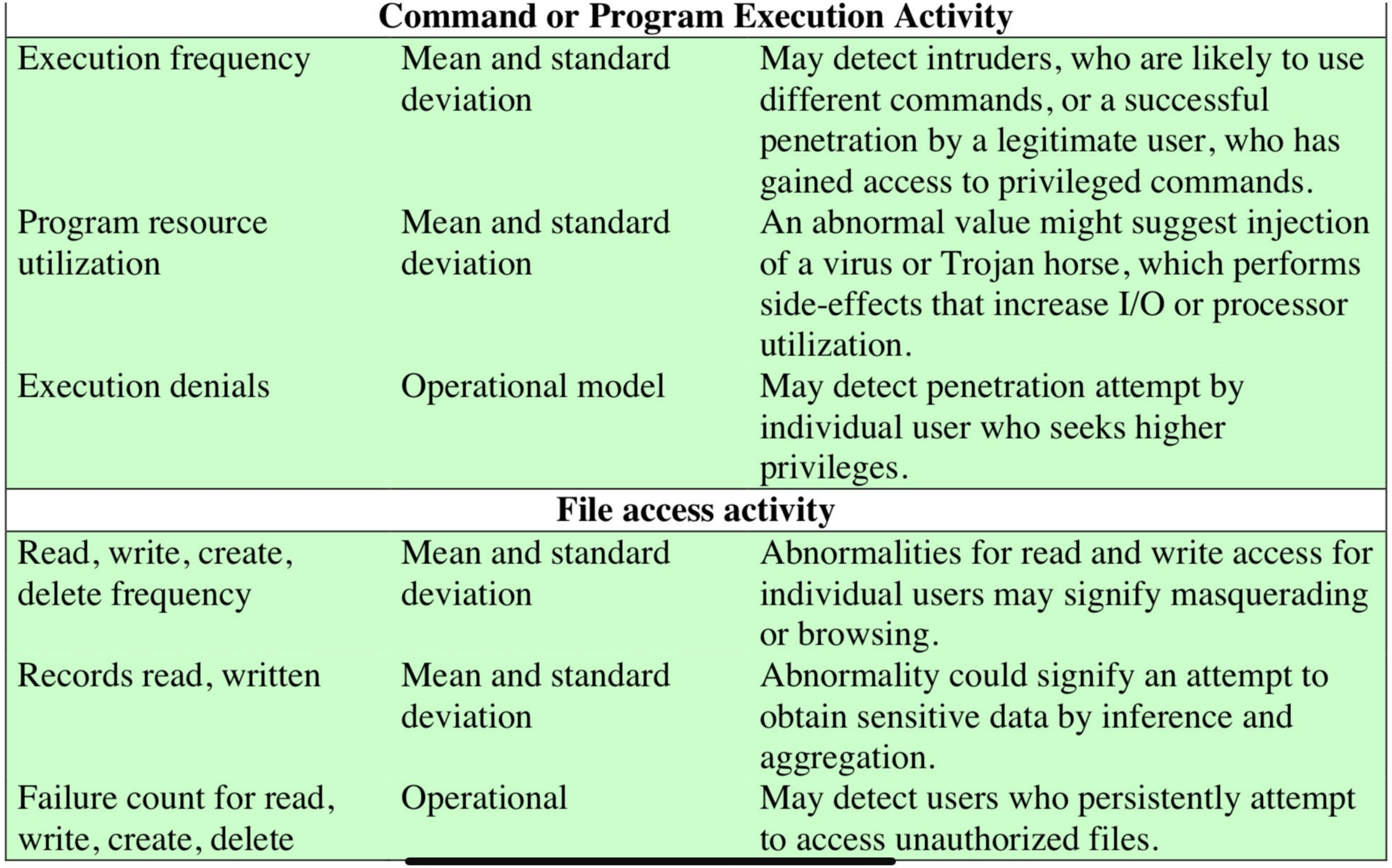
- Given the above Metrics, tests may be carried out to determine whether current activity lies within acceptable limits
  - Mean and Standard Deviation
  - Multivariate Analysis
  - Markov Process
  - Time Series Analysis
  - Operational Analysis
- Main Advantage of Statistical Profiles
  - Prior knowledge of security flaws are not employed
  - The detector program merely looks for deviations from an acceptable norm
  - The approach is not based on system characteristics or vulnerabilities, hence portable as an approach

Solve Model – Metrics and Tests



	Measure	Model	Type of Intrusion Detected
	Login and Session Activity		
1.	Login frequency by day and time	Mean and standard deviation	Intruders may be likely to log in during off-hours.
2.	Frequency of login at different locations	Mean and standard deviation	Intruders may log in from a location that a particular user rarely or never uses.
3	Time since last login	Operational	Break-in on a "dead" account.
4	Elapsed time per session	Mean and standard deviation	Significant deviations might indicate masquerader.
	Quantity of output to location	Mean and standard deviation	Excessive amounts of data transmitted to remote locations could signify leakage of sensitive data.
	Session resource utilization	Mean and standard deviation	Unusual processor or I/O levels could signal an intruder.
	Password failures at login	Operational	Attempted break-in by password guessing.
	Failures to login from specified terminals	Operational	Attempted break-in.

Solve Model – Metrics and Tests



Command or Program Execution Activity		
Execution frequency	Mean and standard deviation	May detect intruders, who are likely to use different commands, or a successful penetration by a legitimate user, who has gained access to privileged commands.
Program resource utilization	Mean and standard deviation	An abnormal value might suggest injection of a virus or Trojan horse, which performs side-effects that increase I/O or processor utilization.
Execution denials	Operational model	May detect penetration attempt by individual user who seeks higher privileges.
File access activity		
Read, write, create, delete frequency	Mean and standard deviation	Abnormalities for read and write access for individual users may signify masquerading or browsing.
Records read, written	Mean and standard deviation	Abnormality could signify an attempt to obtain sensitive data by inference and aggregation.
Failure count for read, write, create, delete	Operational	May detect users who persistently attempt to access unauthorized files.

Set Up and Solve Model: Rule Based Detection

Set up Model – Rule Based Detection

- Method
  - Observe events in system
  - Apply set of rules to determine whether activity is suspicious
  - In general terms we can characterise all approaches as focusing on *anomaly detection* or *penetration identification* (though there exists an overlap between the two approaches)

Set up Model – Rule Based Detection

- Rule-Based Anomaly Detection
  - Similar in approach and strengths to *statistical anomaly detection*
  - In rule based approach
    - Analyse historical audit records
      - To establish patterns of use
      - To automatically generate rules that describe patterns of usage
    - Rules may represent past behaviour patterns for
      - Users
      - Programs
      - Privaleges
      - Time slots
      - Terminals

Solve Model – Rule Based Detection

- Rule-Based Anomaly Detection
  - As with *statistical anomaly detection*, Rule-Based AnomalyDetection
  - Doesn’t require knowledge of security vulnerabilities within system
  - Scheme based on
    - past behaviour
    - Assumption that the future will be similar to the past
    - A large database of rules
      - See e.g.[5] (database contains between 10^4 and 10^6 rules)

Set up Model – Rule Based Detection

- Rule-Based Penetration Identification
    - Based on Expert System
    - Use of rules to identify
        - known penetrations
        - Penetrations that could lead to exploitation of known weaknesses
    - Rules can also be defined
        - To identify suspicious behaviour even when operating within the realms of established usage
    - These are generally
        - Specific to machine and operating system
        - Generated by experts rather than automatically by analysis of audit records

Set up Model – Rule Based Detection

- Rule-Based Penetration Identification
    - Normal procedure
        - Interview system administrators
        - Interview security analysts
        - From these collect a collection of known
            - Penetration scenarios
            - Key events

        that threaten security of target system
    - Approach depends upon skill of those setting up rules

Set up Model– Rule Based Detection

- Rule-Based Penetration Identification
  - Example Heuristics
    - Users should not read files in other users' personal directories
    - Users must not write other users' files
    - Users who log in after hours often access the same files they used earlier
    - Users do not generally open disc devices directly but rely on higher level operating system utilities
    - Users should not be logged in more than once to the same system
    - Users do not make copies of system programs

Set up and Solve Model – Rule Based Detection

- Rule-Based Penetration Identification
    - Typically
        - Audit records are examined as they are generated and compared to the rule base
        - Matches result in an increase in the users *suspicion rating*
        - Once the *suspicion rating* passes a threshold a report of the anomaly is generated
    - State transition models may also be developed
        - http://www.cs.ucsb.edu/research/tech_reports/

Compare with Reality: The Base Rate Fallacy

- We would like our intrusion detection scheme to be efficient at detecting intrusions whilst only infrequently getting it wrong (if at all)
- However due to the nature of probabilistic models it is difficult to attain a high level of detection together with a low level of false alarms
- In general if the incidence of intrusion is low in comparison with the number of legitimate users then the number of false alarms will be high *unless the test is very discriminating*
- The base-rate fallacy had still not yet been overcome in intrusion detection systems