Intrusion Detection - Introduction and Outline

Prof. Dr.- Ing. Firoz Kaderali

Dipl.-Ing. Alex Essoh

Talk Outline

- **Motivation**
- Intrusions
- Intrusion Prevention
- Intrusion Detection
- Major Intrusion Detection Approaches

What is an intrusion?

- A set of actions that attempts to compromise the integrity, confidentiality, or availability of computer resources by causing a DoS, creating a backdoor (Trojan Horse), planting viruses and exploiting software vulnerabilities [AND80].

- An intrusion is a violation of the security policy of a system [KUM95].

- An intrusion is unauthorized access to, and/or activities in, an information system [NST97].

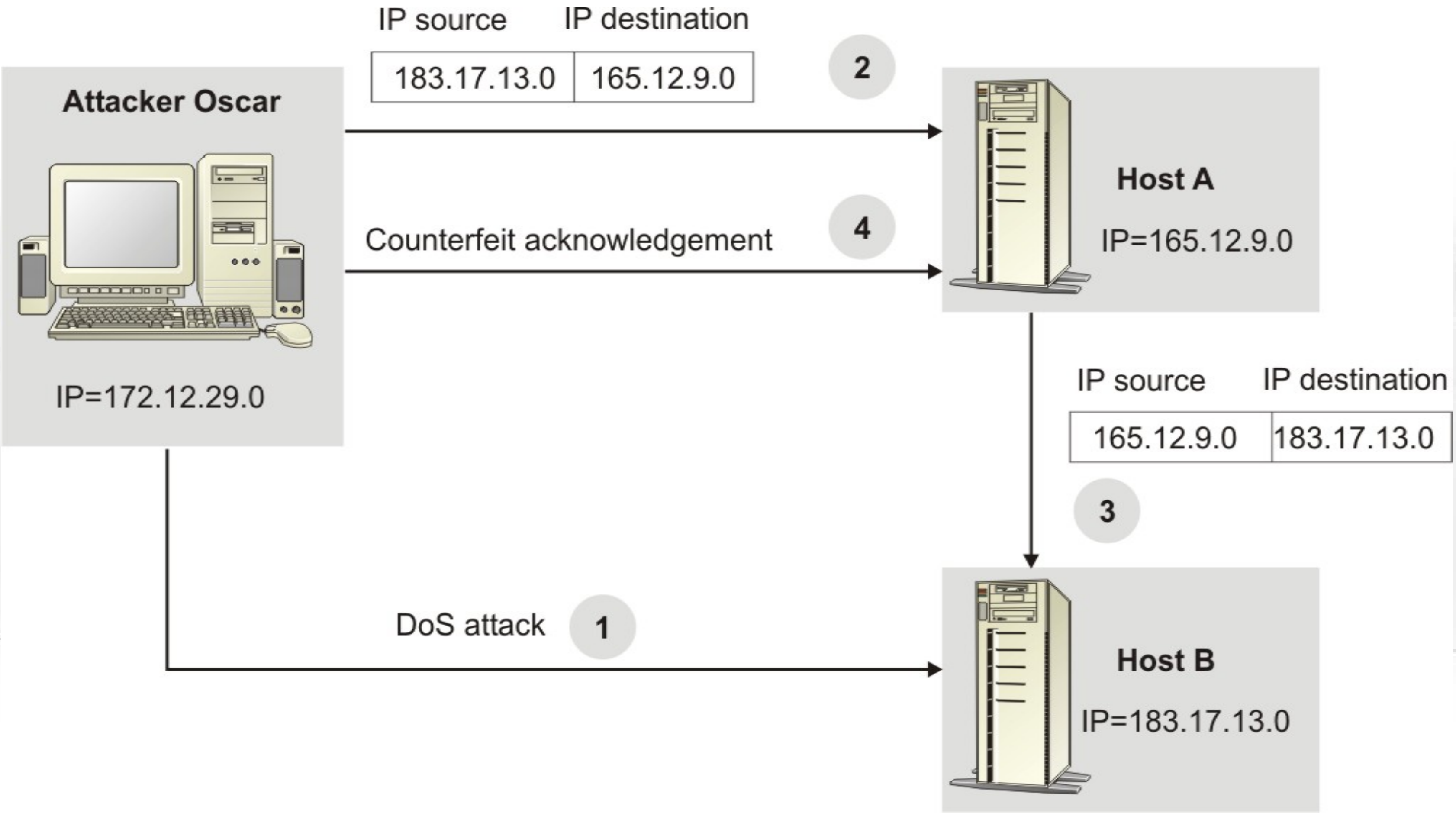
Motivation

- Dramatic increase in incidents
- Attacks are becoming more and more complex and attackers focus on new vulnerabilities
- The resulting damages have been enormous

Talk Outline

- Motivation
- **Intrusion Categories**
  - Protocol related attacks
  - Remote access attacks
  - Malware
  - Denial of Service (DoS)
- Intrusion Prevention
- Intrusion Detection
- Major Intrusion Detection Approaches

IP Spoofing


IP source
IP destination
183.17.13.0
165.12.9.0
2
Attacker Oscar
Host A
IP=165.12.9.0
4
Counterfeit acknowledgement
IP=172.12.29.0
IP source
IP destination
165.12.9.0
183.17.13.0
3
DoS attack
1
Host B
IP=183.17.13.0

Talk Outline

- Motivation
- Intrusions
- **Intrusion Prevention**
- Intrusion Detection
- Major Intrusion Detection Approaches

Firewalls + Access Control

- Firewalls
  - packet filters
  - simple proxies
  - generic proxies implement different applications, but are not able to analyse or filter data streams.

Firewalls are mainly used to filter external traffic, but according to several studies nearly 70-80% of all intruders are internal!

- Access Control
  - Which subject is allowed to access which object.
  - Many attacks are not detectable e.g. user impersonalisation.

Talk Outline

- Motivation
- Intrusions
- Intrusion Prevention
- **Intrusion Detection**
  - Definition
  - Intrusion Detection Types
- Major Intrusion Detection Approaches

Intrusion Detection - Definition

- Intrusion detection is the process of identifying and responding to malicious activity targeted at computing and networking resources [AMO99].
- The process of identifying that an intrusion has been attempted, is occurring, or has occurred [NST97].

Host-based Intrusion Detection Systems (HIDSs)

How are intrusions detected on a host?

- **System Integrity Verification (SIV)**
  - Snapshot of the system (baseline)
  - Cryptographic Check Sums
  - Comparison current state and baseline
  - Example: Tripwire (see http://www.tripwire.org)

- **Automated log files analysis**
  - on each operating system diverse log files are available
  - Windows (Application logs, System logs, Security logs)
  - Solaris Basic Security Modul (BSM)
  - Linux (Last log)
  - Application logs (Web Server)
  - Example: Logsurfer (see http://www.cert.dfn.de/eng/logsurf/)

Web Server logs

- access log

```
192.176.12.173  [06/Nov/2003:10:39:16 +0100]   GET /middle.html HTTP/1.1 200
192.176.12.12   [06/Nov/2003:10:35:14 +0100]   GET /alice.gif HTTP/1.1 304
192.176.12.11   [06/Nov/2003:10:36:14 +0100]   GET /home/user/down.html HTTP/1.1 200
192.176.12.15   [06/Nov/2003:10:37:14 +0100]   GET /print.gif HTTP/1.1 200
192.176.12.9    [06/Nov/2003:10:38:14 +0100]   GET /logo.jpg HTTP/1.1 2000
62.104.86.112   [02/Feb/2004:10:41:19 +0100]   GET /scripts/.\%252e/.\%252e\/winnt/system32/cmd.exe?/c+dir+c: HTTP/1.1\ 404
151.198.253.35  [02/Feb/2004:13:01:46 +0100]    GET /scripts/  .\%\%..\%255c\%255c../winnt/system32/cmd.exe?/c+dir" 404
211.81.24.3     [03/Feb/2004:07:20:34 +0100]    CONNECT 1.3.3.7:1337 HTTP/1.0" 404
```

- error log

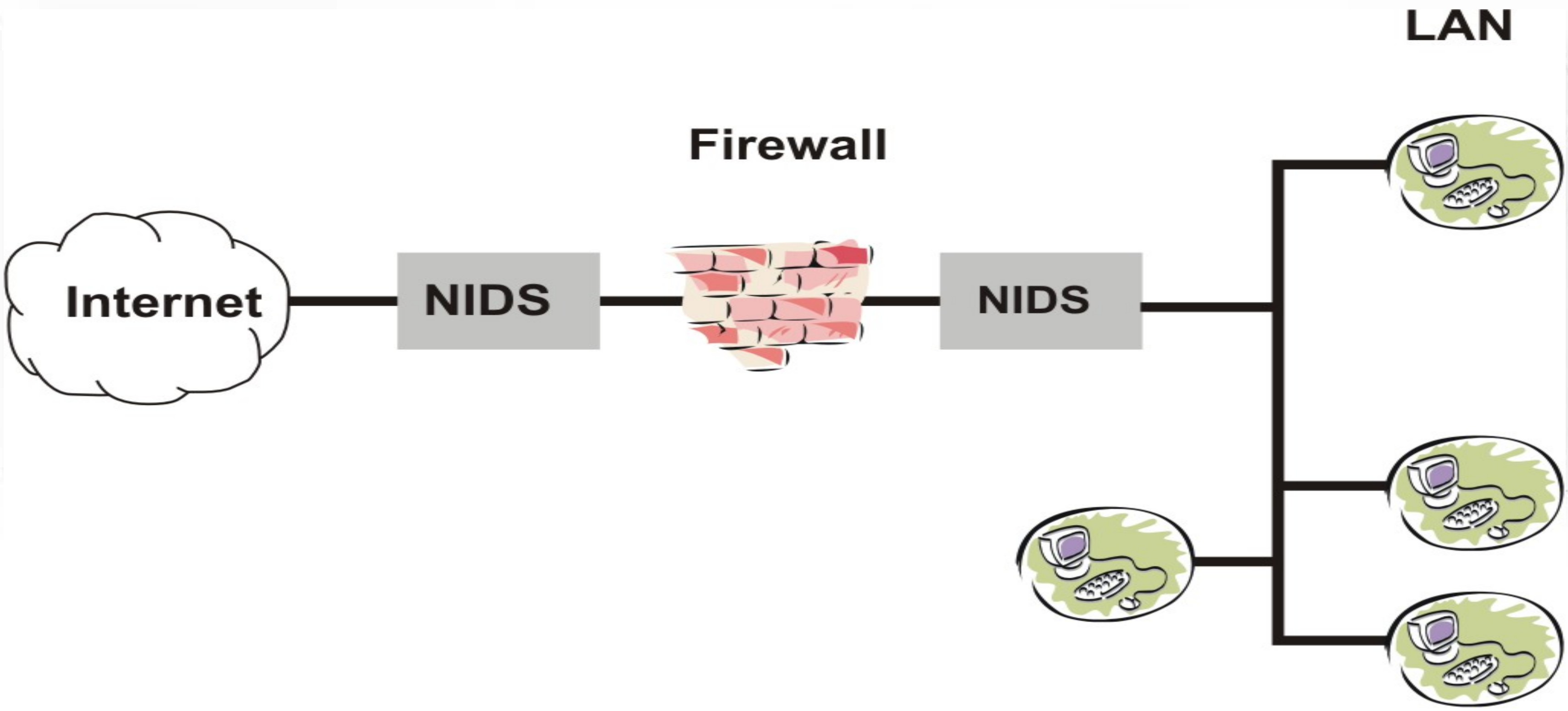
```
217.238.141.213  [12/Feb/2004:09:04:13 +0100] GET /main.php HTTP/1.0 404
217.238.141.213  [12/Feb/2004:09:04:13 +0100] GET /phpinfo.php HTTP/1.0 404
217.238.141.213  [12/Feb/2004:09:04:13 +0100] GET /test.php HTTP/1.0 404
217.238.141.213  [12/Feb/2004:09:04:14 +0100] GET /index.php3 HTTP/1.0 404
128.206.132.141  [12/Feb/2004:10:12:16 +0100] GET /scripts/..\%255c\%255c../winnt/system32/cmd.exe?/c+dir"
218.61.34.188    [14/Feb/2004:14:41:42 +0100] GET /d/winnt/system32/cmd.exe?/c+dir HTTP/1.0 404
```

- last log (Linux)

```
Sep 16 11:55:40   server Failed password for root from 192.176.12.129 port 1137 ssh2
Sep 16 11:55:46  server Failed password for root from 192.176.12.129 port 1137 ssh2
Sep 16 11:55:55  server Failed password for root from 192.176.12.129 port 1137 ssh2
Sep 16 11:55:59  server Accepted password for root from 192.176.9.129 port 1137 ssh2
```

Network-based Intrusion Detection Systems (NIDSs)

- How are intrusions detected on a network?
- NIDS Deployment



Talk Outline

- Motivation
- Intrusions
- Intrusion Prevention
- Intrusion Detection
- **Major Intrusion Detection Approaches**
  - Anomaly Detection
  - Misuse Detection

Anomaly Detection

- Normal behaviour of a subject e.g. a user or a program is profiled (long term behaviour).
- The profile is then compared to the actual behaviour (short term behaviour).
- A new observation is then classified as an anomaly if it does not fit into a predefined tolerance bound.

- **Anomaly Detection Models**
  - Statistical models
    - Markov chains
    - Multivariate analysis

How are anomalies detected using Markov chains?

- The normal user behaviour is fully characterised by
  - the transition probability matrix P

$$P = \begin{bmatrix} p_{11} & p_{12} & p_{13} & \cdots & p_{1n} \\ p_{21} & p_{22} & p_{23} & \cdots & p_{2n} \\ p_{31} & p_{32} & p_{33} & \cdots & p_{3n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ p_{n1} & p_{n2} & p_{n3} & \cdots & p_{nn} \end{bmatrix} \qquad P_{ij} = \frac{n_{ij}}{k} \qquad \sum_j P_{ij} = 1$$

  - and the initial probability distribution Q

$$Q = \left(q_0, q_1, \cdots \cdots q_n\right) \qquad q_i = \frac{k}{N}$$

How are anomalies detected using Markov chains?

- Whenever a sequence of events $S_{1,} S_{2,\cdots,} S_N$ takes place a check is made to see whether this sequence is abnormal or not. The probability of the event sequence occurring is calculated using the Markov chain.

- The probability that a sequence of state $S_{1,} S_{2,\ldots,} S_N$ occurs is [NON00]:

$$P(S_{1,} S_{2,\cdots,} S_N) = q_{S_1} \prod_{i=2}^{N} P_{S_{i-1} - S_i}$$

- A low probability for the sequence transition is likely to be an anomaly.

How are anomalies detected using multivariate analysis?

- Generally we have n previous observations (norm profile) $x_1 \cdots x_n$ and the goal is to check whether a new observation x_{n+1} is abnormal with respect to previous observations [DEN87].

- If the variable to be analysed is a multivariable, the n previous observations can be represented as follows:

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} \qquad X = (y_1, y_2, \cdots, y_p)$$

- The goal is to check whether the new observation $(x_{(n+1)1}, x_{(n+1)2}, x_{(n+1)3,} \cdots x_{(n+1)p})$ is abnormal or not.

- Models (Hotelling´s Test, Chi-Square Multivariate Test) [NON01][NQC01][NON02]

Hotelling´s Test for Anomaly Detection

- Introduced by Harald Hotelling and works as follows:
  - At first the normal behaviour of the multivariable has to be determined by calculating

1. the mean of the multivariable $\overline{X} = \left(\overline{y_1}, \overline{y_2}, \cdots \overline{y_p}\right)$
2. and the variance co-variance matrix $S = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})(X_i - \overline{X})^t$

X_i is the ith observation of the multivariable

  - Then secondly the Hotelling´s test for a new observation $(x_{(n+1)1,} x_{(n+1)2,} x_{(n+1)3,\cdots} x_{(n+1)p})^t$ at discrete time interval (n+1) is calculated:

$$T^2 = (X_{(n+1)} - \overline{X})^t S^{-1} (X_{(n+1)} - \overline{X})$$

Hotelling´s Test for Anomaly Detection

- Thirdly the Hotelling´s test has to be transformed into a F distribution with p and (n-p) degrees by multiplying T² with $\frac{n(n-p)}{p(n+1)(n-1)}$.

- Fourthly the obtained value $T^2 \frac{n(n-p)}{p(n+1)(n-1)}$ is then compared to the tabulated value for a given level of significance α .

If the computed value is greater than the tabulated value, the observation can be classed as abnormal.

Chi-Square Test for Anomaly Detection

- The Chi-Square test is used to see how much the result of the experiment differs from the empirical result.
- The Chi-Square test is calculated as follows [NQC01]:

$$X^2 = \frac{\sum_{i=1}^{p} (X_i - E_i)^2}{E_i} = \frac{\sum_{i=1}^{p} (x_{(n+1)i} - \overline{y}_i)^2}{\overline{y}_i} =$$

- Anomalies can be detected if X² is bigger than the expected tolerance bound.

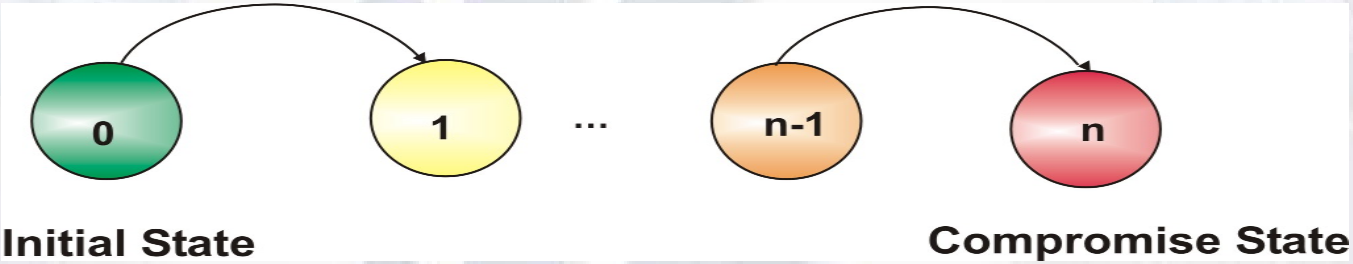
Methods Comparison

- A Markov chain performs better than a Hotelling´s - or a Chi-Square test [NON01].

- Hotelling´s test (95 % detection rate at 0 % false alarm)
- Chi-Square test (60% detection rate at 0% false alarm)
- But after 5% false alarm, Chi-Square performs better.
- Conclusion: the difference between the Hotelling´s - and the Chi-Square test is not very big. [NON02]

Misuse Detection

- Diverse approaches are used:
  - State transition analysis [KOR95]



  - Rule-based misuse detection
    - IF, THEN rules are defined and applied to the data stream.
    - Example: SNORT (see http://www.snort.org)

SNORT

- Snort offers three operational modes:
  - Sniffer
  - Packet logger
  - Intrusion detection
- Each packet traversing the network is analysed in two steps:
  - The Header of the snort rule is applied to the packet, if there is a match
  - options are applied to the rest of the packets.

SNORT Rules

- **Example 1**
  - Land attack: attacker sends an IP packet where the sender IP address equals to the receiver IP address.
  - SNORT rule to detect a land attack:

    <alert ip any any → any any (msg: “DoS Land attack”;sameip;)>

- **Example 2**
  - Unauthorized directory traversal attack: attacker tries to execute malicious commands on a vulnerable Internet Information Server (IIS).
  - SNORT rule to detect an unauthorized directory traversal attack:

    <alert tcp $External_Network any → $My_Webserver $Ports (msg:cmd32.exe access”; content “cmd32.exe”; nocase;)>

Comparison

- Anomaly vs. Misuse

	Anomaly Detection	Misuse Detection
Advantages	Novel attacks can be detected	Lower false positive rate
Disadvantages	Higher false positive rate	Novel attacks can not be detected Database of attacks has to be regularly updated

- Products:
  - RealSecure (see http://www.iss.net)
  - Next-Generation Intrusion Detection Expert System (NIDES) (see http://www.sdl.sri.com/projects/nides/)

Open Issues

- New Methods to significantly decrease the number of false alarms.
- Which variables are suitable for anomaly detection?
- New algorithms to reduce the amount of audit trails for efficient intrusion detection.
- Intrusion correlations.

Literature

- [AMO99] Edward G. Amoroso, Intrusion Detection: An Introduction to Internet Surveillance, Correlation, Trace Back, Traps, and Response, Intrusions.Net Books, 1999

- [AND80] Anderson, J. P., *Computer Security Threat Monitoring and Surveillance*, James P. Anderson Co., Fort Washington, 1980

- [DEN87] Dorothy E. Denning, An Intrusion Detection Model, IEEE Transactions on Software Engineering, volume 13, number 2, pages 222-332, 1987

- [ECK03] Claudia Eckert, IT Sicherheit-Konzepte-Verfahren-Protokolle, 2. Auflage, Oldenbourg, 2003

- [KOR95] Koral Ilgun and R. A. Kemmerer and Phillip A. Porras, State Transition Analysis: A Rule-Based Intrusion Detection Approach, IEEE Trans. Softw. Eng., volume = 21, number = 3,1995, pages = 181--199,

- [KUM95] S. Kumar, *Classification and Detection of Computer Intrusions*, PhD thesis, Dept. of Computer Science, Purdue University, August 1995

Literature

- [MAR01] David J. Marchette, Computer Intrusion Detection and Network Monitoring: A Statistical Viewpoint, Spinger, 2001
- [MUR90] Murray R. Spiegel, Statistik, 2. überarbeitete Auflage, McGraw- Hill Book Company Europe, 1990
- [NON00] Nong Ye, A Markov Chain Model of Temporal Behavior for Anomaly Detection, Proceedings of the 2000 IEEE, June 2000
- [NON01] Nong Ye et al., Probabilistic Techniques for Intrusion Detection based on Computer Audit Data, IEEE Transactions on Systems, MAN, and Cybernetics - Part A: Systems and Humans volume=31, number=4, pages=266--274, July 2001
- [NON02] Nong Ye et all, Multivariate Statistical Analysis of Audit Trails for Host-Based Intrusion Detection, IEEE Transactions on COMPUTERS, volume 51, number 7, pages = 810--820, July 2002

Literature

- [NQC01] Nong Ye and Qiang Chen, An Anomaly Detection Technique based on a Chi-Square Statistic for detecting intrusions into information systems, Quality and Reliability Engineering International, volume=17, pages =105--112,
- [NST97] NSTAC Intrusion Detection Subgroup Report, Dec. 1997
- [RIC03] Robert Richardson, Eight AnnualCSI/FBI Computer Crime and Security Survey, 2003 http://www.visionael.com/products/security_audit/FBI_CSI_2003.pdf
- [SAC93] Lothar Sachs, Statistiche Methoden, Planung und Auswertung, Springer Verlag, 1993
- [SPE03] Ralph Spenneberg, Intrusion Detection für Linux-Server, Markt + Technik, 2003
- [SYM03] Symantec Internet Security Threat Report 2003

End

Thanks for your attention.

Prof. Dr. Firoz Kaderali
Dept. Communication Systems
FernUniversität Hagen
email: firoz.kaderali@fernuni-hagen.de