Click here to view PowerPoint presentation; Press Esc to exit

Intrusion Detection Introduction and Generics

H.A.M. Luiijf, R. Coolen
TNO Physics and Electronics Laboratory
P.O. Box 96864
2509 JG The Hague, The Netherlands

E-mail: *luiijf@fel.tno.nl, coolen@fel.tno.nl*

Abstract

This paper gives an introduction to intrusion detection systems (IDSs) in order for the general audience to understand the specific R&D aspects discussed by the other symposium papers. The generic model of an IDS presented in this paper gives a common viewpoint for the study and discussion of functionality and components of IDS in a product independent way.
Furthermore, the necessity for real-time intrusion detection is highlighted using Logicon's time-axis model of an attack.

Introduction

Increasingly, NATO Forces operate in multinational coalitions and connect the NATO (nation) networks to non-NATO nations (NNNs), non-NATO international organisations (NNIOs), and non-governmental organisations (NGOs). Increasingly, operational requirements demand the sharing of information and integrating communication and information systems (CISs) of NATO nations and/or other coalition Forces. The use of interconnected modern information and communication technologies enhance the situational awareness and the strive for information dominance.
Internal and external threats to CIS, amplified by interconnecting with CIS of other nations and organisations, require early and often real-time warnings about intrusions and other irregularities in the NATO CIS as well as effective counter-measures. This to reduce the risks associated with potential unauthorised access to, compromise of, and control over NATO information and that of its members.
Intrusion Detection Systems (IDSs) are technical means that focus on the detection type of measures against intrusions in and to a CIS.

The Task Group on Information Assurance under the Information Systems Technology Panel of the NATO Research & Technology Organisation (NATO/RTO/IST/TG03) investigated the generics of Intrusion Detection Systems in 2000 – 2001 and published a report [9]. This paper highlights the main models and issues from that report as a foundation for this Real-time Intrusion Detection symposium.

Terms and definitions

In literature, different terms and definitions are used for intrusion and intrusion detection. In this report definitions are chosen in accordance with [1] and [2], but with a focus on the military operational CIS environment:

Attack: A deliberate intrusion in a CIS.

Attacker: The person, group, organisation or state that performs an attack.

Defender: The person, group or organisation (i.e. NATO agency or NATO nation) that is responsible for the CIS to be protected.

Intruder: The person, group, organisation or state responsible for an intrusion.

Intrusion: A deliberate or accidental unauthorised access to, activity against, and/or activity in, a CIS.

Intrusion Detection: The process of identifying that an intrusion has been attempted, will occur, is occurring, or has occurred.

Target: The CIS that an intrusion is aimed at.

Paper presented at the RTO IST Symposium on "Real Time Intrusion Detection", held in Estoril, Portugal from 27-28 May 2002, and published in RTO-MP-101.

Note that the definition of an intrusion includes intrusions that have an intentional or unintentional intend, harmful or harmless consequences, and concern both intrusions by insiders and outsiders (by definition respectively affiliated to the defender organisation or not).

Both unintentional intrusions and attacks can result in damage. This damage can concern the availability, integrity and/or confidentiality of the CIS.

Security measures

In order to see where IDSs fit in to the overall range of security measures, intrusion detection is positioned as one of the coherent security measures against an incident (e.g. an intrusion). The measures are presented using the security incident cycle, which is visualised in figure 1.

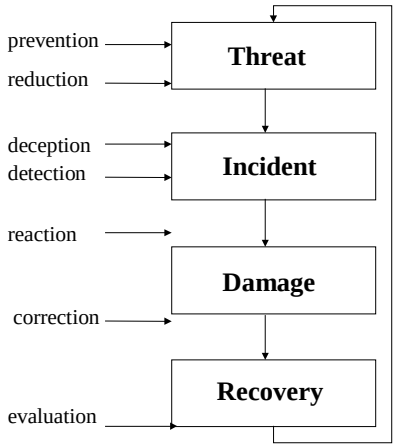

Figure 1*:* Security incident cycle

The incident part of the cycle consists of four elements: the threat, the incident, the occurrence of damage, and the recovery. The following different types of security measures are related to these elements: prevention, reduction, deception, detection, reaction, correction, and evaluation.

The security incident cycle has to deal with threats to the confidentiality, integrity and availability of the CIS. A defender first of all takes *prevention measures*. These measures prevent a threat from becoming a reality. An example of a prevention measure to protect an internal network is a Boundary Protection Device (BPD). Another example of a prevention measure is scanning for known vulnerabilities in a CIS and thereafter correcting these vulnerabilities by implementing patches or changing configuration parameters.
Reduction measures are measures that are performed in advance to reduce possible damage of an incident such as an intrusion. Examples of reduction measures are redundant systems, limitation of bandwidth, and regular back-ups.
Deception measures are a special type of security measures. They have the purpose to give false information to intruders, to reduce the possibility of an incident, to allow easier detection of an incident, to slow intruders down, or to obtain operational benefits over the intruding party.

Prevention, reduction and deception measures reduce the probability and the impact of an incident. However, this does not exclude possible occurrence. Therefore, the defender takes *detection measures*.
All intrusions have to be detected as early as possible. In this way, the defender does not loose valuable time over the intruder. This time can be used to identify the intruder and to take more extensive prevention, reduction, and deception measures to minimise damage and to maximise the possibility for a proper reaction. Intrusion detection is the main focus of IDSs. However, there is a tendency that other security measures such as reaction and deception are also incorporated in IDS-products.

After an intrusion is detected, the defender takes *reaction measures*. These reaction measures can be repressive in order to block the repetition of the intrusion. The reaction measures can also include tracing an intruder. Furthermore, if the operational authority for the CIS decides to start a process to press charges

against an intruder/attacker, inforensic[1] evidence often needs to be collected between the moment of the first intrusion related events, a successful intrusion in, and the recovery of the CIS [2].

When an intrusion results in damage to the integrity or availability of information, the next step in the security incident cycle is to take *correction measures* to undo at best the damage that was done. The operational status of vital parts of the CIS has to be reconstituted as soon as possible. This is where *reduction measures* such as back-ups prove their usefulness.

The final step in the security-incident cycle consists of an effectiveness *evaluation* of the security measures taken. Questions might be: 'what went well and what went wrong?' And what lessons can be learned and how to prevent a reoccurrence of the intrusion in the future?

Note that prevention, reduction, and detection measures should be designed according to the defence-in-depth principle [3]. That is the attacker or intruder should have to overcome multiple lines of defence before he/she is able to breach the confidentiality, availability, or integrity of the NATO, NATO member's or coalition partner CIS(s).

Generic IDS Model

The generic IDS model from [9] is visualised in figure 2. The purpose of this model is to function as a common viewpoint for the study and discussion of functionalities and components of IDS in a product independent way. Different components are distinguished and described in a logical order. Furthermore some additional definitions are provided.

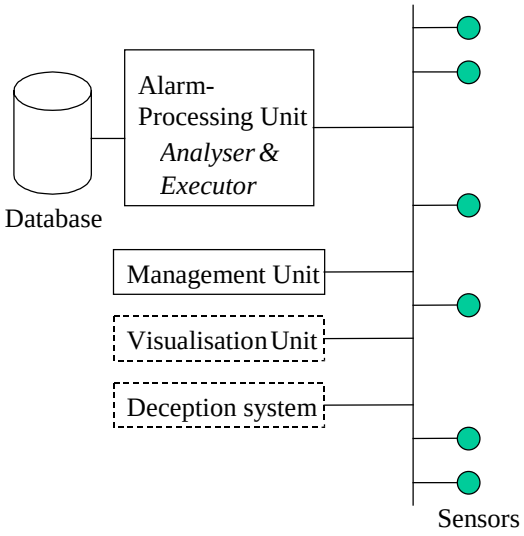

Figure 2: IDS Generic model

Sensors

The description of the generic components of an IDS is started with introducing *sensors.* These are the generic components of an IDS that collect *activity*. This activity can include network traffic, user misbehaviour, application misbehaviour and so forth. In the intrusion detection community it is common to distinguish between two types of activity: network activity and host activity.

Network activity: the activity present at the network is the network traffic, which can be categorised in:

- Low level protocols (ISO layers 2 though 4 e.g. TCP, UDP);
- Application and service level protocols (e.g. SMTP, HTTP, FTP);
- Content (of e.g. e-mail or web pages).

---

1 Information forensics, reconstruction and recovery: the application of forensic techniques to investigate crimes involving, either directly or indirectly, information and communication technology (ICT).

2 A workshop on Inforensics will be organised by NATO/RTO/IST/RTG-003 in October 2002.

Host activity: at the hosts (including clients, servers and routers) several forms of activity are present, caused by:
- Users: the person operating on a host, e.g. identified by a login account;
- Systems: hardware, operating system;
- Network services (e.g. PKI, DNS);
- Applications: e-mail, web browsers, and so forth.

Examples of sensors are network interfaces in promiscuous mode and tools that read log-files. An IDS can have multiple sensors. Based on the type of activity the sensors collect, the following classification of IDSs can be made:
1. *Host-based IDS (HIDS)*: the IDS looks at activity on a host;
2. *Network-based IDS (NIDS)*: the IDS looks at the network traffic either in (near) real-time or inspects via log-files at regular intervals;
3. *Hybrid* IDS: the IDS has sensors collecting host and network activity.

Alarm-processing unit
The alarm-processing unit is the generic component of an IDS that pre-processes and analyses the activity collected by the sensors. Furthermore, the alarm-processing unit controls the reaction to be taken by the IDS in reaction of a detected intrusion. The processes of analysis (by the analyser) and reaction (by the executor) are essential for an IDS and are described in more detail.

There exist two main classes of alarm-processing units based on the technique they use for analysing activity to detect intrusions [1] misuse and anomaly detection. These classes of IDS analyser techniques (misuse and anomaly detection) will be discussed below. But first, the difficulty of analysing activity with respect to false alarms is discussed.

No alarm-processing unit is infallible in analysing potential intrusion related activity. The alarm-processing unit in an analyser may fail to detect intrusions, or sound the alarm when no intrusion has occurred. Four cases in the operation of an analyser are distinguished (Table 1).

Table 1: Four cases in the operation of an event analysis.

	Intrusion	**No Intrusion**
IDS Alarm	An intrusion has occurred, and the IDS generated an alarm *(correct alarm)*	No intrusion has occurred, but the IDS has (erroneously) detected an intrusion *(false alarm)*
IDS Rejection	An intrusion has occurred, but the IDS has not generated an alarm *(false rejection)*	No intrusion has occurred and the IDS has not detected an intrusion *(correct rejection)*

Related to these four cases two parameters of an IDS are defined:
- The *accuracy* of an IDS: the number of *correct alarms* divided by the number of *correct alarms plus false alarms*; is a parameter for the relative number of correct alarms. The more accurate an IDS is, the fewer false alarms it generates and the higher this parameter is.
- The *completeness* of an IDS: the number of *correct alarms* divided by the number of *correct alarms plus false rejections*; is a parameter for the relative number of correct alarms. The more complete an IDS is, the fewer intrusions remain undetected and the higher this parameter is.

In the ideal case, an IDS would be 100% complete (it detects all intrusions) and 100% accurate (it produces no false alarms). However, detecting an intrusion is a very difficult task. This comes partly forth from the base-rate fallacy problem described in [4]. The base-rate fallacy problem shows the need for very accurate IDSs. If

an IDS generates too many false alerts, the operating and managing personnel will have no confidence in the system!

Analyser techniques
Analysers that use the *misuse detection* method operate by searching for very explicit activity and/or patterns of activity. Misuse detection is also called detection-by-appearance.

A number of known *intrusion patterns* (also known as intrusion *signatures* or *rules*), that specify the features, conditions, arrangements and interrelationships among activity that leads to break-in or other misuse are stored beforehand in the IDSs knowledge database. The IDS collects activity and looks whether one of the stored intrusion patterns occurs. If an intrusion pattern is detected, the IDS will generate an alarm.
A detection-by-appearance IDS can only detect *known* intrusions, but once it detects an intrusion, it can usually specify exactly how the intrusion has occurred.
In a sense the misuse detection concept is paradoxical, because the intrusions have to be known beforehand. One could argue that a CIS should not be vulnerable to known intrusions. However, in practice it can be infeasible to remove all vulnerabilities from a system, because this is costs to much resources. Therefore the misuse detection IDS can certainly provide an important role to detect known intrusions, intrusion attempts and other intrusion related activity.

An advantage of a misuse-detection IDS is that it is not only useful to detect intrusions, but that it will also detect intrusion attempts; a partial signature may indicate an intrusion attempt. Furthermore, the misuse-detection IDS could detect port-scans and other events that possibly precede an intrusion.

A disadvantage of a misuse-detection IDS is that only known intrusions are detected. No protection is offered against novel attacks, or new variants of existing intrusions. More crucially, a small variation in the form/structure of an attack can invalidate a signature.

Misuse detection is the most used technique in current NIDSs. There are two different types of NIDS, smart and raw, depending on whether they look at patterns in low-level protocol activity or application-level protocol activity [5].

Smart: IDSs that have logic implemented that understands the target protocol. They will parse the request and perform (optimised) signature matching based on known rules pertaining to the protocol. They will attempt to behave like a real web server would behave, at the expense of additional code and slowness.

Raw: Also referred to as 'packet grep' style IDSs, they typically just scan the unprocessed raw data for key strings. The benefit of this is speed only. The term raw is not used in a derogatory manner, but rather to identify that these IDSs usually deal with the raw data directly, rather than interpreting the protocols.

Event analysers that use the *anomaly detection* method operate by examining the behaviour of the activity. Anomaly detection is also known as detection-by-behaviour. The ‘normal’ behaviour pattern of activity is stored beforehand in the IDS. The current activity is then continuously collected in real-time and analysed to see whether its behaviour significantly deviates from the stored behaviour. The IDS signals significant deviation between these behaviours as a (possible) intrusion.

An anomaly detection IDS is confronted by two related problems:
1. *Description problem*: How to describe the behaviour of activity in an effective and efficient manner?
2. *Comparison problem*: Given a stored behaviour pattern of activity and a current behaviour, when do these two deviate enough to constitute a possible intrusion?

An IDS based on anomaly detection is classified according to how it deals with these two problems, see [7].
The main problem with anomaly-detection IDSs is that it is hard to describe the ‘normal’ behaviour of activity, because of e.g. the unpredictable behaviour of the end-users. This results in IDSs generating a lot of false alarms. This is the main reason that anomaly detection IDSs are hardly used in practice and are mostly at a research stage.
The main advantage of a good working (few false alarms) anomaly detection IDS would be that in a sense ‘unknown’ intrusions can be detected. The systems however detect only the fact *that* an intrusion has occurred

rather than *how* it has occurred. This creates a situation where the IDS does not need a priori knowledge of specific security flaws in a CIS.

In [4], a strict anomaly detection model is described which has the distinct feature that it generates no false alarms by definition. The proposed strict anomaly detection model is as follows. An IDS should use precise definitions of 'use' for activity in a CIS, in accordance with security policy. Any deviation of these definitions is a security policy violation. In case of the normal anomaly detection the 'use' behaviour is not strictly defined but merely a description of normal behaviour of activity. The key-advantage of the strict anomaly detection model is that new attacks can be detected, while no false alarms are generated.

Human interaction with the analyser can be an important aspect of analysing activity. Therefore the alarm-processing unit should be able to cope with human intervention in the decision process of whether activity indicates an intrusion.

A characteristic of an IDS is the frequency of the analysis. Three categories are distinguished:
1. *Continuously*: events are collected and thereafter analysed, as they occur, - often in real-time.
2. *Periodically*: events are collected and analysed periodically from the subject. An example is system log files that are analysed every hour.
3. *Initiated under special circumstances*: e.g. when the system administrator suspects an intrusion.

The information about known intrusions and/or the normal behaviour of the activity is stored in a *knowledge database*. The alarm-processing unit can also store (information about) collected activity in a *storage database* that can be of interest in the future for example in a digital forensics investigation.

Reaction
When the IDS decides that certain activity indicates an intrusion, an alarm is generated by the executor component. This alarm can either be *passive* or *active:*

Passive: an IDS generates an alarm, which can be a log file message, a pop-up screen, a pager message and so on, or a combination thereof.

Active: the generic model includes IDSs that have (optional) active components that can generate automatic reactive control signals. These control signals could for example tighten a BPD, increase the IDS' sensitivity, shutdown a connection, divert network traffic to a decoy system or shut down hosts that are under attack.

An IDS can operate in close co-operation with network management systems. Alarms can be incorporated in network management systems. Active IDS-alarm components might send network management control messages to different components of the network.
In general a reaction of an IDS will be closely related to (response) management of the system. Although a reaction can be partly automated, in practice human interaction and decision making will be needed to come to a balanced reaction.

Management
The management of IDSs is crucial for efficient and good deployment of IDSs in military (and also government or corporate) environment. The management of an IDS is divided in four categories [1]:
- Detection management
- Response management.
- Update management.
- Availability management.

Besides the issues of scalability and protection of management operations is described.

Detection management involves communicating with the IDSs via e.g. a graphical user interface that visualises possible intrusions. Furthermore it can involve manual analysis of data by a manager, e.g. to double-check IDS alarms.

The intrusion detection system, when it is signature or rule based, has to be updated very regularly. As we already noted, new attacks arise every moment, hence the updating of signature based IDSs is an ongoing task.

The process of updating the system is called update management. Commercial IDSs do not have a continuous 24h updating process of signatures.

Once an intrusion is confirmed, responses have to be managed. Most of the actions that have to be taken can not be performed automatically by the executor-part of an IDS, but require manual intervention

Availability management deals with ensuring that the system is available at all times. Both the hardware and software components can go out of service and then need maintenance. Furthermore, IDSs can be under a denial-of-service attack.

Central management is an important property for IDSs in a CIS. Especially when scalability of the intrusion detection capabilities are concerned. Maintenance and updating of different systems becomes difficult when network environments are growing and IDSs are not centrally managed. As there is no standard management protocol for IDSs, IDSs from different manufacturers can not yet be managed from a single central management system.
Management operations themselves must be protected from intruders as well. For example, if management commands travel over an in-band network, there is a risk that the IDSs themselves get compromised, this includes the vulnerability to denial-of-service attacks. This issue becomes especially important when IDSs are centrally managed. Authentication and confidentiality can be provided by using encryption. In the mean time, most commercial IDSs already have encrypted communication implemented.

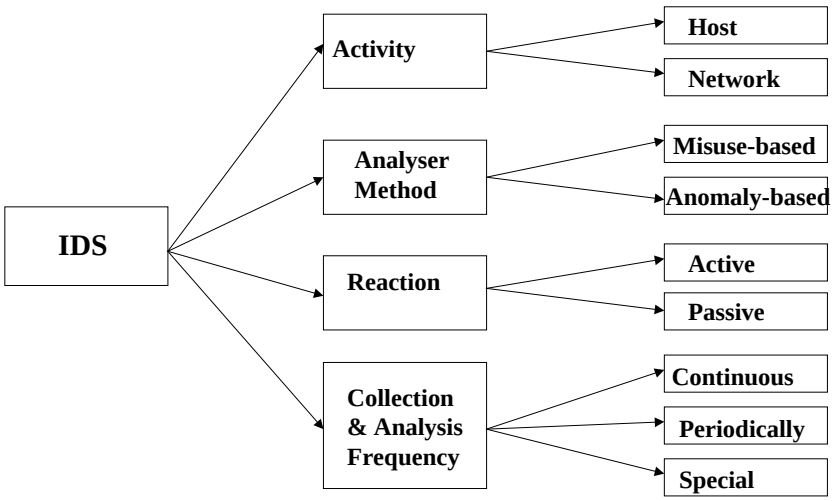

Figure 3: IDS characteristics

Visualisation
Closely related to the management of an IDS is the optional visualisation of intrusions. In order to obtain informational situation awareness, visualisation of intrusions in a CIS is an important issue. This is incorporated in the generic models for Intrusion Detection by the optional intrusion visualisation unit. It can combine information of intrusions with other information such as the network (link) performance (e.g. availability and usage). The visualisation component can be tightly linked to an IDS alarm processor, but preferably also able to fuse information from various types of sources, such as network performance monitors.

Deception systems
As an optional generic component an IDS can contain a deception system. These can be used to attract attackers to certain parts of the infrastructure. Preferably the deception system is on a stand-alone system. This has the advantage, that all network traffic directed to the system is suspicious and indicates an intrusion. In this way, an intrusion can be detected at an early stage. The entire intrusion can be recorded and hence novel intrusion techniques can be learned. Moreover, when the attacker spends time intruding the deception system, valuable time is gained over the attacker. This time can be used to protect the real CIS and/or to trace the intruder.

Deception systems, however, have disadvantages as well. Firstly, when compromised they can be used as a stepping stone to further compromise the CIS. Building the system in a virtual machine (jail) can make this a lot harder for the attacker. Secondly, deception systems add complexity to the CIS. This may lead to increased vulnerabilities. Finally, deception systems have to be managed at the cost of resources.

Distributed hierarchy of IDSs

Like the figure above of the IDS generic model, Figure 4 visualises a distributed CIS where each sensor consists of an IDS operating at a lower level of abstraction in the network. This distributed design is often seen in commercial products.
In a similar way, different alarm messages from a distributed IDS corresponding to different CISs can be communicated with an IDS one level of abstraction higher, by adding an extra layer to the model.
For instance, this could be a NATO-wide IDS where for instance different NATO member states send intrusion alarm messages to a NATO-wide intrusion visualisation and management facility. This allows management of intrusion detection at NATO level.

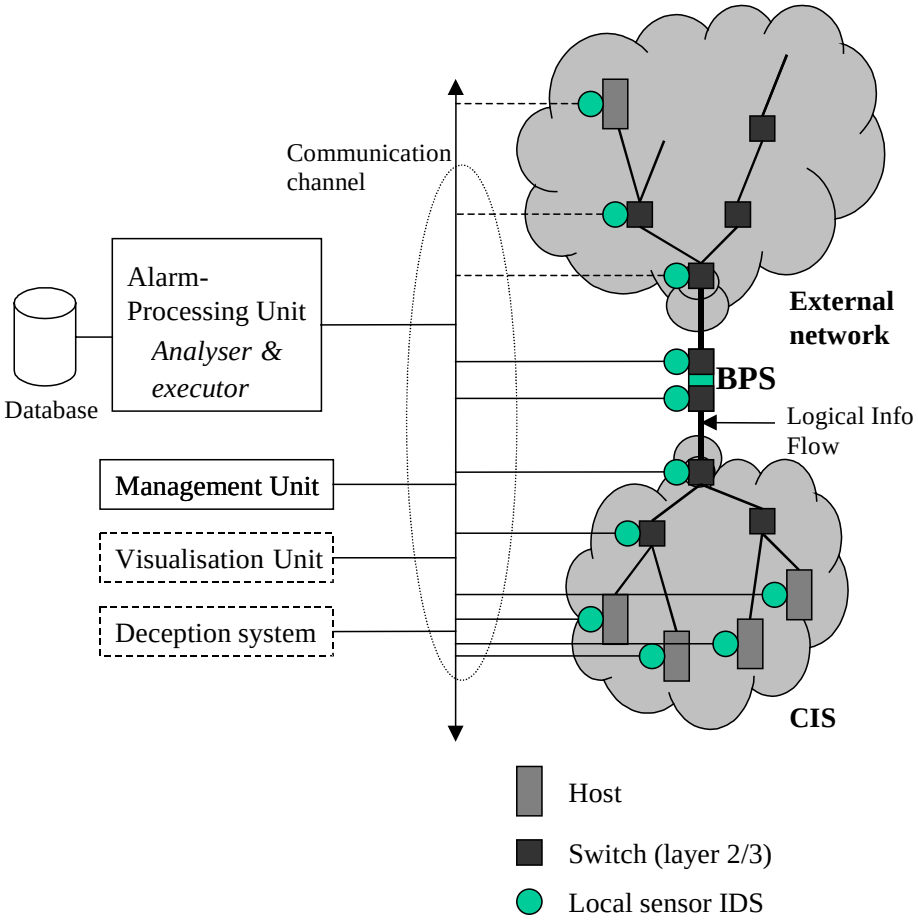

Figure 4: Distributed IDS

Another example is an IDS in a NATO military operation, where CISs of different NATO members, NNNs, NGIOs and NGOs release (parts of) their local intrusion information to the higher IDS-management and visual layer. This in order to acquire an overall operational awareness of enemy information warfare activities.

Real-time detection, a necessity

Before discussing the impact of the real-time necessity on IDSs, the Logicon Inc. time-axis model is used to describe the time-line of an attack (~ a deliberate intrusion) [6]. This attack scheme is presented in Figure 5.

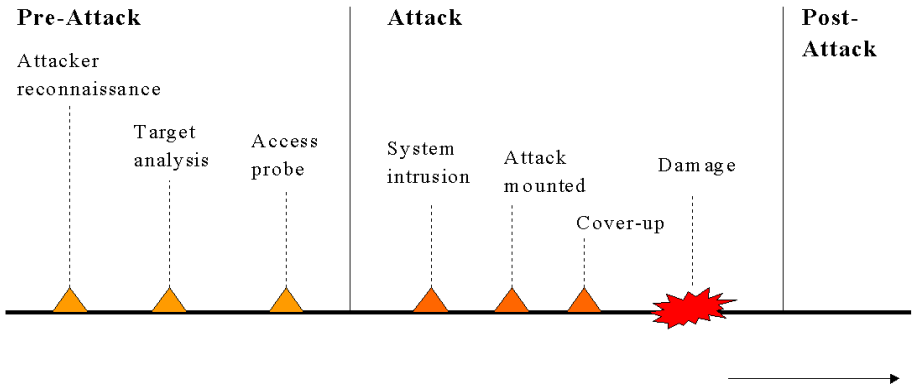

Figure 5: Time axis model of an attack (source: [6])

The actual attack is preceded by a pre-attack stage. The attacker will begin this stage by defining an end-state with regard to the CIS (~ target). This end-state is a clearly defined and obtainable objective. Desired results may be denial-of-service, acquisition of sensitive information, and/or establish and maintain access to the CIS.

After setting the objective the attacker will seek to identify and define problems associated with breaching the target defences, gather information and make assumptions about the CIS, develop possible courses of action (COA), and analyse each COA. In the time-axis model, three steps are distinguished in the pre-attack stage:

1. *attacker reconnaissance,*
2. *target analysis,*
3. *access probing.*

In the first step: *attacker reconnaissance*, the attacker starts acquiring critical information about the CIS. This includes execution of most, if not all, of the following steps: foot printing, scanning, enumeration, vulnerability mapping, and social engineering (i.e. using social skills to obtain info from e.g. employees). The second step: *target-analysis* consists of analysing the available information, making assumptions and then developing multiple COAs. In the third step: *access probing*, the attacker tests the COA, and then selects the best COA. The testing is often done, by sending probes to the CIS or by stimulating the CIS. For a “complex” attack, the pre-attack stage can last a long time. In the following stage of the time-axis model, the actual attack starts. Hereafter, the attacker will try to cover up the operation and/or leave a backdoor (e.g. a Trojan Horse or a kernel patch) in the system. An attack can be very hard to recognise when the cover up operation is performed well. After the damage is done, the post-attack stage starts. This is the stage, where the defender will try to take corrective measures and so forth.

The time-axis model can be related to the incident cycle described above. Both models have a point where an intrusion leads to damage. Where the incident-cycle models the defender’s actions in the periods of time before and after the damage, the time-axis model shows the attacker’s actions for these two periods of time. The point of detection of the incident from the incident cycle can also be related to the time-axis model. When an attack is detected in the pre-attack stage, this is called *pre-attack detection*. Similarly when an attack is detected in the actual attack stage this is called *attack-detection*. It is also important to recognise that a system was attacked and that possible damage has occurred or that there is a security breach. The detection of an attack in the post-attack stage is defined as *damage detection* or *post-attack detection*. Has crucial information been modified? Is there a backdoor present in the system?

It is crucial for defence organisations, but also for most other organisations, that intrusions and attacks are detected as fast as possible. It is impossible to take reaction, correction and evaluation measures if it is unknown that an intrusion will occur, is occurring or has been occurred. Hence, damage can not be controlled and minimised, which might result in financial disasters, or worse.

For intrusion detection this implies that IDSs should operate in a real-time manner. Real-time intrusion detection can be seen from two viewpoints. Firstly, within the boundaries of technology an alarm should be available to the response managers as soon as possible. And secondly, an intrusion should be detected as early as possible on the timeline of an attack. An important property of the analyser to achieve the latter is the ability to correlate data.

Correlation of data

At most stages of the attack, there is activity in the (target) CIS, which can be detected. The process of interpreting, combining and analysing the information of all available sources (such as IDS in the target CIS, but also available information from other sources) is called *correlation* or *fusion*.

Different sensor IDSs (located in the CIS or even in external network) can collect information from different stages of the attack. To optimally use this information for early warning the IDS should be able to correlate the information in real-time. This should especially include information from the pre-attack stage, since the first signs of an attack are visible in this stage. The information used could be *in-band information* or *all-band information* [8]. In-band information is all information from activity inherent to the target system. All-band information can be any other information that can be used in the correlation, including information from human intelligence sources.

Conclusion

The generic IDS model, the different aspects of IDS and all issues mentioned above, shall give the general audience enough background for understanding the issues addressed by the other practical and research papers later during this symposium. For a more extensive introduction to the generics of IDS, the reader is referred to [9].

References

1. 'ISO/IEC PDTR 15947, Information technology – Security techniques – IT intrusion detection framework', ISO/IEC JTC 1/SC 27 N2691.
2. The President's National Security telecommunications Advisory Committee, Network group Intrusion Detection, 'Subgroup Report on the NS/EP Implications of Intrusion Detection Research and Development', December 1997.
3. 'Technical and Implementation Directive for the Interconnection of Communication and Information Systems (CIS)', INFOSEC.
4. Sasha, Beetle, 'A strict anomaly detection model for IDS', Phrack 56, source: www.phrack.com., 6-11-2000.
5. Rain Forest Puppy, 'A look at Whisker's anti-IDS techniques', 1999, source: www.wiretrip.net/rfp/, 6-11-2000.
6. Paul Zavidniak, Logicon Inc.: 'Achieving Information Resiliency in the Defence Environment', Information and Security and Data Security Congress, February 2000.
7. Esmaili, M., R. Safavi-Naini and J. Pieprzyk, 'Intrusion detection: a survey'. In: S.J. Chung (ed.). Proceedings of the 12th International conference on Computer Communications 21-24 August 1995 in Seoul. Amsterdam, 1995, pp. 409-414.
8. Amoroso, 'Intrusion Detection: an introduction to Internet surveillance, correlation, traps, trace-back, and response', intrusion.net books, 1999.
9. Coolen, R., Luiijf, H.A.M. (2001), *Intrusion detection, Generics and State-of-the-Art*. NATO RTO/IST TR-49, NATO RTA, Paris, France. On-line: www.rto.nato.int.