A new taxonomy for intrusion detection

C.J.Tucker[1,2], S.M.Furnell[1], B.V.Ghita[1] and P.J.Brooke[3]

[1]Network Research Group, University of Plymouth, Plymouth, United Kingdom
[2]Stochastic Systems Limited, St Austell, Cornwall, United Kingdom
[3]School of Computing, University of Teesside, Middlesbrough, United Kingdom
e-mail: intrusion@stochastic.co.uk

Abstract

A new taxonomy is proposed in which the type of output and the data scale over which an intrusion system operates is used for classification. This taxonomy allows a graphical comparison of different intrusion systems to be undertaken in terms of their footprint on an intrusion matrix. It is proposed that quantitative comparison of systems can only be undertaken at points of overlap of their footprints and that overlap specific measures are needed for this comparison. New areas of application for intrusion systems are also discussed.

Keywords

Computer Security, Intrusion Detection, Taxonomy

1. Introduction

Intrusion detection has a long history, dating back to the work of Anderson (Anderson, 1980). Since then, various analysis techniques, ranging from support vector machines, through to data mining and expert systems, have been used as part of the detection engine (Mukkamala and Sung, 2003). Many complete systems have been constructed and operated on live computer systems (eg (Allen et al., 2005)). Despite over 25 years of research, the topic is still active, in part due to the rapid development of information processing systems and their vulnerabilities, but also due to fundamental difficulties in achieving an accurate declaration of an intrusion. Intrusion systems are noted for high false alarm rates and considerable research effort is still concentrated on finding effective intrusion, non-intrusion discriminants.

This paper proposes a new taxonomy which aims to improve the comparison of intrusion systems. The proposed taxonomy considers the different type of outputs that can be produced by intrusion systems, along with the type of information used to determine the intrusion, as the basis for their comparison.

2. Background

A number of taxonomies for intrusion detection have already been proposed. One of the earliest was undertaken by Debar and classified intrusion systems according to their detection method, behaviour on detection, audit source location, or usage frequency (Debar et al., 1999). This was later extended to include the detection paradigm, as either state- or transition-based (Debar et al., 2000). Axelsson offered an alternative taxonomy in terms of the detection

principle and operational aspects, such as whether operation is continuous or in batch mode (Axelsson, 2000).

Each of these taxonomies provides insight into the operation of intrusion systems and is a useful framework for identifying new research opportunities, but they are not a good basis for their comparison as they are based on the internal properties of such systems. A taxonomy based on the applicability of intrusion systems is a more fundamental comparison approach as it describes their use, rather than the details of their implementation.

Consider, for example, two network intrusion systems. System A is misuse-based whilst System B is anomaly-based. During a series of intrusion events both these systems will indicate the intrusion state of the network segment they are monitoring, possibly to different degrees of accuracy (i.e. their detection and false alarm rates). Much work has been published on the quantitative comparison of such systems (e.g. (Abouzakhar and Manson, 2004)), using analysis techniques such as receiver operating characteristic (ROC) or lift curves. However, in addition to the presence of an intrusion, System A will often indicate the type of attack and the exploit being used, on the basis of the specific signatures that are triggered. Comparing System A with System B via a ROC curve or confusion matrix will not include this important property of System A and thus is not a fair comparison method.

3. A new intrusion taxonomy

The taxonomy proposed in this paper is inspired by the work of Johnson in the formation of images (Johnson, 1958). He studied the ability of human operators to find and correctly classify objects within complex images. The objects were relatively small and thus the a priori probability that a specific area of the image contained an object was very low, a situation analogous to intrusion events within a background of normal network or host activity (Axelsson, 1999). Johnson defined the following types of operator tasks (amongst others):

- **Detection** – the ability to say that something of interest is present in the image.
- **Recognition** – the ability to determine the class of object present, such as a car or aircraft.
- **Identification** – the ability to determine the type of object present, such as the make of car or the type of aircraft.

The most important aspect of Johnson's work was the definition of minimum criteria necessary for successful completion of the above tasks. Using similar task definitions as a starting point, this study proposes to divide the output of intrusion systems into one of 5 categories, with the first 3 loosely in line with Johnson, as follows:

- **Detection** – in which the system outputs an indication of a state change within a network or host. There is no classification of the nature of the change, apart for the assumption that this indicates the occurrence of a possible intrusion. The principal use of such systems is for data rate reduction so that other systems (either automated or human) can investigate further.
- **Recognition** – in which the intrusion systems are capable of declaring the type of attack, such as Distributed Denial of Service (DDoS), reconnaissance, or User to Root (U2R).

- **Identification** – in which the system is capable of declaring the exploits used to achieve the intrusion, such as buffer overflow or an application-specific vulnerability.
- **Confirmation** – in which the attack plan is deduced, allowing attack-specific countermeasures to be deployed rather than coarse measures, such as disconnection of the internet access or isolation of key business servers.
- **Prosecution** – in which evidential quality data is generated identifying the originator of the intrusion.

As an example of the use of this taxonomy consider a simple anomaly intrusion system comparing the utilised network bandwidth with historical values. Such a system would be categorised as an intrusion detection system. It would be able to declare that something unusual is happening within the network but declaring with any certainty that the anomaly was caused by an intruder is not likely to be achievable.

As another example consider a Snort intrusion system operating on a single network segment (Roesch, 1999). When a rule is triggered and an alert declared, there is considerable attack-related information available. Often, rules are created to alert when the signatures of specific attacks are present. Thus, when such a rule has been triggered, the intrusion system can identify the exploit being used. In this respect Snort is acting as an intrusion identification system.

In addition to considering the output from an intrusion system, further insight can be achieved from an analysis of the data scale over which the system is operating. In modern network systems four data scales can be considered:

- **File** – monitoring the status of individual files for unauthorised access or change.
- **Host** –monitoring the applications running on and the behaviour of an individual host.
- **Network** – monitoring the packets exchanged between hosts, servers and other network devices to assert the presence of an intrusion.
- **Enterprise** – monitoring traffic originating from trusted sources of an organisation which operate in the presence of other, less trusted data sources.

The File, Host and Network data scales have been used in other studies (e.g. (Bace and Mell, 2001)). The separation of Network data scale into two sections, as introduced by this paper, is believed to be a novel concept. The principal difference between the Network and Enterprise data scales is the mixing of trusted and untrusted data streams within the same network segment. This is most often encountered in virtual private networks (VPN) between an office location of an organisation and its remote staff or trusted partners, via the Internet. VPNs are separated from the untrusted data streams using encryption schemes and well-known protocols. However, this separation may become subject to the same technology, policy or configuration vulnerabilities as other parts of the information processing system. Therefore, it is likely that an individual responsible for a secure network would want to know that their VPN communications were subject to intrusion attempts. Intrusion systems therefore need to extend their data scale applicability to include the Enterprise. This is a technically challenging problem.

4. The application of the taxonomy

This taxonomy can be applied in a number of ways. The remainder of this paper will examine its use to create an intrusion footprint on a grid or matrix formed from the output type and data scale elements of the taxonomy. The use of this footprint for comparison of systems will then be shown.

4.1 Intrusion matrix

The combination of intrusion output type and data scale can be shown as an intrusion matrix, as in figure 1.

	FILES	HOST	NETWORK	ENTERPRISE
PROSECUTION	Protective Monitoring, Secure Data Vaults	Protective Monitoring, Secure Data Vault	Anti-spoofing, Trust Management	-
CONFIRMATION	A Priori Assessment	A Priori Assessment	AI Techniques	-
IDENTIFICATION	Malware Signatures	Malware Signatures	Exploit Signatures	-
RECOGNITION	File Hashes	Malware Signatures, Resource Anomalies	Traffic, Protocol or User Anomalies	-
DETECTION	File Hashes	Sys Calls, Registry Use, Resource Anomalies	Traffic, Protocol or User Anomalies	Quantum Cryptography

Figure 1: Intrusion System Taxonomy Matrix

Also shown in figure 1 are some of the techniques that can be applied within a particular output type and data scale. For example, malware signatures or resource anomalies can be used in intrusion recognition systems operating at the Host data scale. Much of this matrix is covered with techniques that have been extensively studied. Of particular note is the absence of techniques at the Enterprise data scale.

4.2 Intrusion system footprint

The intrusion matrix can be used to plot a footprint for different intrusion systems. The footprints are determined from an analysis of the intrusion system outputs to determine which of the 5 output categories the system is capable of producing and what data scale is used to create the output. For example, figure 2 shows the footprints of a number of different intrusion paradigms. Figure 2a shows the footprint of a representative anti-virus software (AVS) package. They typically include both virus-specific signatures and heuristics that respond to anomalous behaviours. This means they operate from the Detection to the Identification output types. Since the attack plan can often be determined by reverse

engineering of the virus, AVS packages can also be considered to operate at the Confirmation output type.

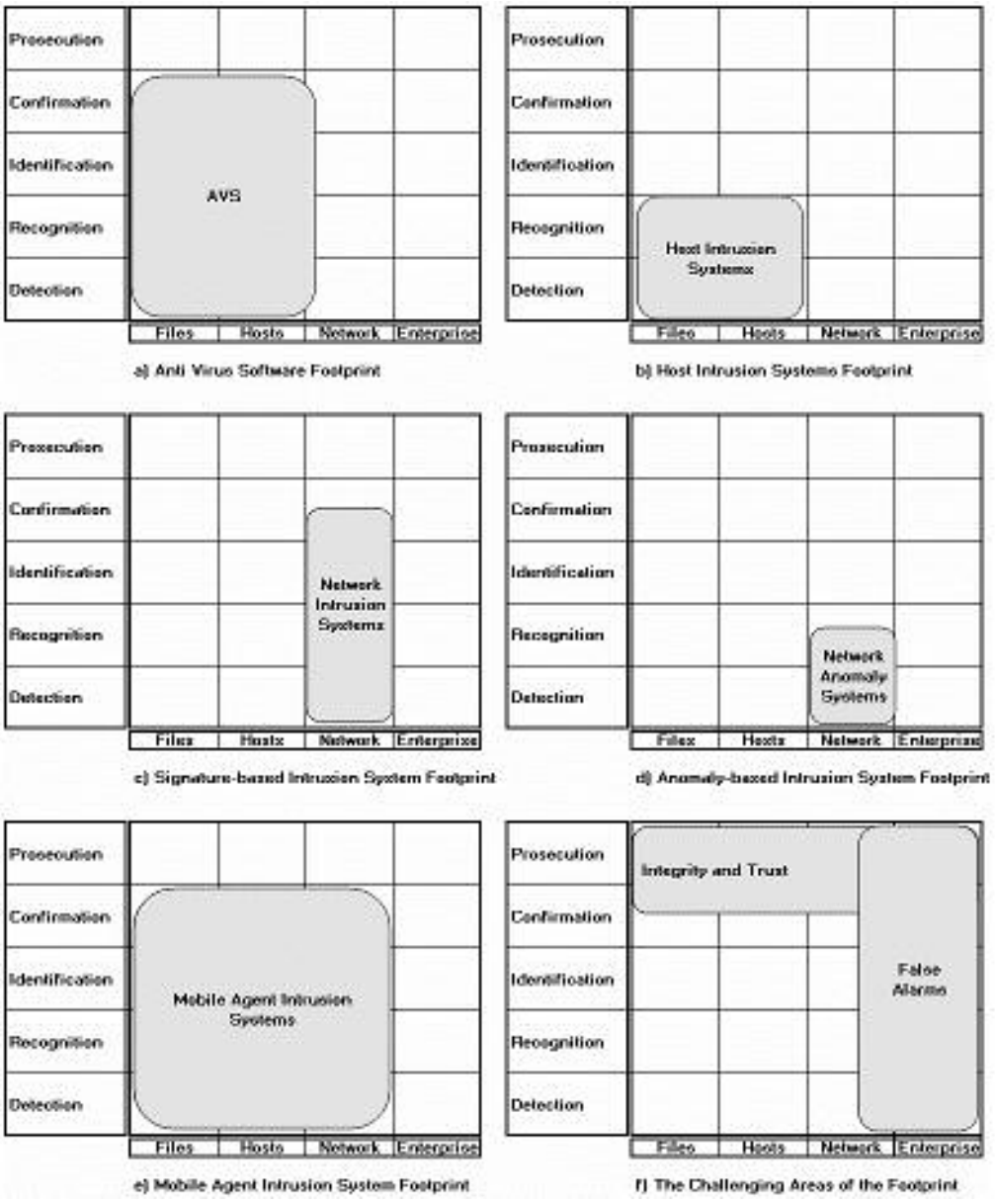

Figure 2: Intrusion System Footprints

A footprint of a host-based intrusion system is shown in figure 2b. To create this footprint it was assumed that anomaly techniques are applied and therefore the intrusion system is only capable of Detection or Recognition. Confirmation, or the determination of the specific exploit or vulnerability used (Identification), are unlikely to be achievable with confidence when using an anomaly based system. Host-based intrusion systems using signature techniques would be expected to operate at the Identification and Confirmation levels, depending on the discrimination capabilities of the signatures.

Figures 2c and 2d show network-based intrusion systems using signature and anomaly detection respectively. These figures highlight the principal differences to be at the higher output types of Identification and Confirmation. Snort is a typical example of a signature

based intrusion system. On its own it is unable to perform the plan determination required for full Confirmation. However, when multiple Snort sensors are deployed at strategic parts of a network, it is possible to determine an attack plan from the patterns of signatures that are triggered. An additional module would be required to integrate the information and hence determine the plan. Hence, the Confirmation output type is shown partially covered by the footprint.

Figure 2e is the most interesting, and shows the extensive footprint that could be achieved by intrusion systems based on mobile agents. On the assumption that mobile agents could be created to examine the status of files, applications running on a host, and packets on the local network segment, they offer the widest range of data scales of any other technique. Also, their payload could include integrated anomaly and signature based techniques, and when combined with a communications capability this could give them the potential to provide output types up to Confirmation. It may even be possible that techniques for Enterprise data scales and Prosecution could be integrated as they become available.

Finally, figure 2f shows the current challenges faced by intrusion systems. The Prosecution output type requires high integrity information to be gathered and secured from change. Whilst this is a common requirement in secure systems it must be achieved to the levels necessary to allow criminal prosecution, within a system that has intruders present. For the Enterprise data scale, the technology challenge appears to be the development of discriminants that will separate intrusion and non-intrusion events in mixed-trust data flows.

4.3 Comparison of intrusion systems

The intrusion matrix can be used to provide a comparison between systems. A qualitative comparison can be made by examining the footprint of each system. Large footprints are likely to represent systems which provide a broader range of applicability and a wider range of output information during an intrusion. Small footprints would be typical for systems which are very specific in their application.

A more quantitative comparison can be made by examining the performance of systems where their footprints overlap. Each element of the intrusion matrix is accompanied by a set of performance metrics relevant to the output data type. These performance metrics could include false alarm rates, intrusion probabilities, or confusion matrices measured in such a way as to be appropriate to the position within the intrusion matrix. As an example consider a single element within the intrusion matrix, say the (Network, Identification) element. If the footprint of two intrusion systems overlap on this element then performance metrics relevant to Identification should be calculated for the 2 systems. The probability of identification could be determined as a function of the false alarm rate, to produce Identification ROC curves. Examination of the ROC curves at this overlap point within the intrusion matrix would allow comparison of the systems in the role of intrusion identification. A fair comparison would require the examination of performance metrics at all points of overlap on the intrusion matrix.

Some of the elements of the intrusion matrix presented have been extensively studied and can be considered commercial successes. For example, AVS packages are very successful at providing confident alerts at the Files and Host data scales (Post and Kagan, 1998). Such software can be very specific, identifying the virus and hence, by implication the “plan” of the originator of the virus. Heuristic algorithms can provide a degree of detection capability in

which the AVS indicates that there is a virus present but is not specific about its type. AVS packages are also well known to provide a high alert probability with a low false alarm rate. Thus a large area of this matrix can be achieved with very high performance.

Meanwhile, some of the elements of the intrusion matrix are poorly understood at this time. Effective techniques at the Enterprise data scale are rare and of limited applicability. This applies at any of the intrusion output levels. The trusted data stream may be present with untrusted streams and on untrusted network equipment (e.g. Internet backbone routers). Intrusion systems are unlikely to be able to operate outside of the trusted systems of the enterprise, leaving Enterprise scale intrusion systems to rely on remote diagnosis of intrusion behaviour.

It can therefore be seen that there are 3 aspects of the intrusion matrix that are important in determining the performance of an intrusion system, namely:

a) The number of elements of the matrix that an individual system footprint covers as this can indicate the applicability of the system.
b) The position of the elements of the footprint within the intrusion matrix, as some element positions present an inherently challenge to achieve high performance
c) Only the elements that overlap are of any significance in the quantitative comparison of intrusion system.

5. Conclusions and further work

This paper proposed a novel taxonomy for intrusion systems, based on the type of information the system is capable of providing, as well as the data scale over which it operates. It allows a graphical comparison of systems to be undertaken, using their footprint on an intrusion matrix constructed from the output capabilities and data scale. A more precise, quantitative comparison can then be undertaken at overlapping elements within the footprint, using metrics specific to the type of output. Thus ROC curves based on Probability of Recognition versus Recognition False Alarms would be created to compare systems that have a footprint overlap at the Recognition output type.

Finally, the inclusion of AVS within this taxonomy opens the challenging and interesting alternative of placing intrusion systems on a more theoretical basis. The work of Cohen (Cohen, 1987) has already established theoretical limits on the detectability of viruses, proving that no algorithm can perfectly detect all possible viruses. More recently Li has proposed a theoretical basis for intrusion, but this work has yet to reveal any useful conclusions (Li et al., 2005). It is hoped that this taxonomy will build on this theoretical basis and lead to a better understanding of the limits of performance for intrusion systems, as well as providing an improved framework for their comparison.

6. References

Abouzakhar, N. S. & Manson, G. A. (2004), "Evaluation of intelligent intrusion detection Models", *International Journal of Digital Evidence,* 3.

Allen, W. H., Marin, G. A. & Rivera, L. A. (2005), "Automated detection of malicious reconnaissance to enhance network security", *SoutheastCon, 2005. Proceedings. IEEE*

Anderson, J. (1980), "Computer security, threat monitoring and surveillance", Fort Washington PA, James P Anderson Co.

Axelsson, S. (1999), "Base-rate fallacy and its implications for the difficulty of intrusion detection", *Proceedings of the 1999 6th ACM Conference on Computer and Communications Security (ACM CCS), Nov 2-Nov 4 1999*, Singapore, Singapore

Axelsson, S. (2000), "Intrusion detection systems: A survey and taxonomy", Department of Computer Engineering, Chalmers University.

Bace, R. & Mell, P. (2001), "Intrusion detection systems", *NIST Special Publication on Intrusion Detection System*.

Cohen, F. (1987), "Computer viruses: Theory and experiments", *Computers and Security,* 6**,** 22-35.

Debar, H., Dacier, M. & Wespi, A. (1999), "Towards a taxonomy of intrusion-detection systems", *Computer Networks-the International Journal of Computer and Telecommunications Networking,* 31**,** 805-822.

Debar, H., Dacier, M. & Wespi, A. (2000), "A revised taxonomy for intrusion-detection systems", *Annales Des Telecommunications-Annals of Telecommunications,* 55**,** 361-378.

Johnson, J. (1958), "Analysis of image forming systems", *Proceedings of the Image Intensifier Symposium*, US Army Engineering Research Development Laboratories, Fort Belvoir, USA

Li, Z., Das, A. & Zhou, J. (2005), "Theoretical basis for intrusion detection", *Systems, Man and Cybernetics (SMC) Information Assurance Workshop, 2005. Proceedings from the Sixth Annual IEEE*

Mukkamala, S. & Sung, A. H. (2003), "A comparative study of techniques for intrusion detection", *Tools with Artificial Intelligence, 2003. Proceedings. 15th IEEE International Conference on*,

Post, G. & Kagan, A. (1998), "The use and effectiveness of anti-virus software", *Computers & Security,* 17**,** 589-599.

Roesch, M. (1999), "Snort - Lightweight intrusion detection for networks", *Proceedings of USENIX 13th Systems Administration Conference (LISA '99)*, Berkeley, CA