A RESPONSE-ORIENTED TAXONOMY OF IT SYSTEM INTRUSIONS

M.Papadaki[†], S.M.Furnell[†], B.M.Lines[†] and P.L.Reynolds[‡]

[†] Network Research Group, Department of Communication and Electronic Engineering, University of Plymouth, Plymouth, United Kingdom
[‡] Orange Personal Communications Services Ltd, St James Court, Great Park Road, Bradley Stoke, Bristol, United Kingdom.
e-mail: nrg@plymouth.ac.uk
Web: http://www.plymouth.ac.uk/nrg

KEYWORDS

Intrusion Detection Systems, Intrusion Taxonomy, Intrusion Response.

ABSTRACT

The ability to select and initiate appropriate response(s) is an issue that is often neglected in Intrusion Detection Systems (IDS). In order to address the problem, a means is required to consider different potential security breaches, the differing contexts in which they may occur, and the differing potential consequences. Current intrusion taxonomies have limited application in this regard, considering categories of intrusions that could not be detected by an IDS, or representing potential results in too few dimensions to enable any fine-grain selection of response options. This paper presents an overview of a new taxonomy, which is specifically targeted towards enabling the consideration of responses. A number of generic incident and target categories are identified, encompassing the most common forms of intrusion/attack and the contexts in which they may occur. An assessment of the likely results is then presented in each case, considering the security impacts, the time available to respond, and further potential attacks that may be initiated as a result. By encompassing alternative targets, and considering multi-dimensional results, the taxonomy provides a means of differentiating the incidents on the basis of the responses they require, rather than by characteristics of the attack method or their security impacts alone.

INTRODUCTION

Intrusion Detection has been an active research area within the computer security domain for more than 15 years. The challenges associated with this area have so far been largely concentrated on the process of detecting an intrusion. However, automation of the next stage after detection, the response to an incident, is a significant issue that has not been adequately addressed and therefore requires further research in its own right.

Intrusion response is defined as the process of counteracting the effects of an intrusion. It includes the series of actions taken by an Intrusion Detection System (IDS), following the detection of a security-related event. The justification for advancing the automated response capability of IDS is twofold: firstly, to reduce the significant overhead that manual response poses to the administration of increasingly large and complicated IT infrastructures, and secondly, to cope with the widespread use of automated scripts that can generate attacks of distributed nature.

In order to select appropriate responses, it is necessary to know more than just the type of incident that has occurred, or the basic security impact that has resulted. However, many current intrusion classification taxonomies provide little understanding beyond this level. As such, a new taxonomy has been developed as the basis for studying the issue of response, aiming to consider incidents and identify their different results in different contexts. It is intended that this taxonomy will give insight into the process of selecting appropriate responses and forming the basis of decision-making in an automated responder system (Furnell and Dowland 2000)

The discussion begins by summarizing previous work that has been conducted in relation to intrusion and attack taxonomies, before proceeding to present details of the new approach. The concept of the response-oriented taxonomy builds upon previous ideas, originally introduced by Furnell et al (2001).

CURRENT INTRUSION TAXONOMIES

Previous research has given rise to a number of intrusion taxonomies, each of which presents an alternative view of the situation. Brief summaries of a number of notable approaches are given below.

A common method of classifying security incidents is according to the impacts or outcomes resulting from their occurrence. This has led to a number of result-based taxonomies of incidents and attacks. In such approaches, all attacks are grouped into basic categories according to their result, aiming to give more insight into their severity. An example is a taxonomy devised by Cohen (1995) that

includes result categories such as *Corruption, Leakage*, and *Denial*. *Corruption* is defined as the unauthorised modification of information, *leakage* is when information ends up where it should not be, and *denial* is when computer or network services are not available for use. Another result-based taxonomy is specified by Russell and Gangemi (1991), who define similar outcome categories, but use a different set of terms (i.e. *secrecy* and *confidentiality* instead of *leakage; accuracy, integrity*, and *authenticity* instead of *corruption;* and *availability* instead of *denial*).

Although result-based taxonomies can be useful in providing a meaningful association between different types of attacks, the end result of an attack is not the only significant characteristic and thus it represents only one aspect of the problem. In order to detect, respond, and specify protection, it is necessary to have some classification of the incidents that lead to the results. In this respect, there are also a number of prior works that can be considered.

Cheswick and Bellovin (1994) classify attacks into the seven categories listed below:

- Stealing passwords - methods used to obtain other users' passwords
- Social engineering - talking your way into information that you should not have
- Bugs and backdoors - taking advantage of systems that do not meet their specifications, or replacing software with compromised versions
- Authentication failures2 - defeating of mechanisms used for authentication
- Protocol failures - protocols themselves are improperly designed or implemented
- Information leakage - using systems such as *finger* or the *DNS* to obtain information that is necessary to administrators and the proper operation of the network, but could also be used by attackers
- Denial-of-service - efforts to prevent users from being able to use their systems.

Although this approach provides a general overview, including the main categories of intrusions, it is not specified in any further detail, and thus is too general to provide any insight to the relationship among different classes of attacks or their different characteristics.

Neumann and Parker (1989) developed an intrusion taxonomy based on a large number of incidents reported to the Internet risks forum. The taxonomy classifies intrusions into nine categories, according to key elements that might indicate a particular type of incident. Table 1 below summarises the overall scheme.

NP 1 EXTERNAL MISUSE	Nontechnical, physically separate intrusions
NP 2 HARDWARE MISUSE	Passive or active hardware security problems
NP 3 MASQUERADING	Spoofs and Identity changes
NP 4 SUBSEQUENT MISUSE	Setting up intrusion via plants, bugs
NP 5 CONTROL BYPASS	Going around authorised protections/controls
NP 6 ACTIVE RESOURCE MISUSE	Unauthorised changing of resources
NP 7 PASSIVE RESOURCE MISUSE	Unauthorised reading of resources
NP 8 MISUSE VIA INACTION	Neglect of failure to protect a resource
NP 9 INDIRECT AID	Planning tools for misuse

Table 1: SRI Neumann-Parker taxonomy

An extension of the Neumann-Parker taxonomy was produced by Lindqvist and Jonsson (1997), which further refines security incidents into intrusions, attacks and breaches. It examines these issues from a system-owner perspective, based on a number of laboratory experiments. The results of these experiments indicated a need for further subdivision of the Neumann-Parker classes 5, 6 and 7, as shown in Table 2 below. Their work provides further insight into the process of spotting aspects of system elements that might indicate an intrusion.

Extended NP5 CONTROL BYPASS	Password attacks, spoofing privileged programs, utilizing weak authentication
Extended NP6 ACTIVE RESOURCE MISUSE	Exploitation of write permissions, resource exhaustion
Extended NP7 PASSIVE RESOURCE MISUSE	Manual browsing, automated browsing

Table 2: Lindqvist and Jonssen extension of the Neumann-Parker taxonomy

A final example is provided by Howard (1997), who follows a different approach by focusing on the process of an attack, rather than classification categories. Howard's taxonomy establishes a link through the different potential *attackers* (classified as hackers, spies, terrorists, corporate raiders, professional criminals and vandals) and the *tools* and *access methods* that they may utilise, leading to the *results* that enable the attackers to achieve their objectives. This taxonomy was based on the analysis of real incidents, as reported to the CERT/CC from 1989 to 1995, and thus represents a very valuable tool for systematically studying attacks. Having said this, it does not present a comprehensive top-level classification of

intrusion incidents, or yield an appropriate classification that could be used to determine the required response – a criticism that could also be levelled at the other examples considered here.

Although most of the existing taxonomies succeed in contributing to the systematic study of intrusions, they are not immediately applicable to the domain of automated intrusion detection and response systems. From a detection perspective, it is clear that a number of the incident classifications identified (e.g. social engineering, physical tampering), and issues such as the objectives of attackers, could not be detected or determined by an automated system. In addition, they do not give any insight into the issue of response. A taxonomy that would serve this purpose ought to give consideration to the classification criteria, which will include aspects such as incident type, target, and/or potential impact. This will lead to indication of generic response categories, considering what can be done to halt an attack in progress, reduce its impact and/or prevent reoccurrence. The discussion of such a taxonomy is the focus of the next section.

A RESPONSE-ORIENTED TAXONOMY

The aim of the new taxonomy is to determine the effect an incident has on specific targets, and demonstrate how that may influence the response decision process. In order to demonstrate that concept a set of incidents have been used and are listed below:

1. Information gathering (Probe / Scan, Sniff)
2. Authentication failure (Masquerade / Spoof, Bypass)
3. Software compromise (Buffer Overflow, Flood / Denial of Service (DoS)
4. Malware (Trojan Horse, Virus / Worm)
5. Misuse (Unauthorised Alteration, Unauthorised Access)

As with the previous taxonomies, the selection of incidents is by no means exhaustive, but the five top-level categories aim to encompass the most significant set of incidents that affect current systems. Also, the description of the incidents used in the taxonomy aims to preserve a high level of abstraction, in order to include as many cases of incidents as possible. So, for example, although there are many different methods of launching Denial of Service attacks (such as SYN Flooding, SMURF attacks, Ping of Death, Trin00, and others), their ultimate effect on a system is similar, and it is this that will be the main determinant of the desired response(s). The five incident categories, and example incidents, are described more fully later in this section, following discussion of the other elements of the taxonomy.

Another important characteristic that can influence response is the *Target* of the intrusion, since the same incident can have different impacts upon different targets. The target groups considered in the new taxonomy are as follows:

- *External server:* Public-facing servers that are accessible from external networks and represent the public image of the host organization (e.g. web, email, DNS, FTP servers). Ideally, if configured correctly, external servers should not contain or facilitate access to confidential information, but ought to provide uninterrupted service to clients.
- *Internal server:* A server accessible only within the internal network of the organization (e.g. intranet web and file servers).
- *User workstation:* Computing units used by average users, likely to contain information specific to a particular user and their role within the organisation.
- *Network Component:* Networking equipment such as routers, switches, firewalls, which may be targeted as a means of accessing other systems or subverting operations.

This is by no means a detailed or exhaustive list, but it is sufficient to give a high level abstraction of the different elements that might be targeted in a typical organisation.

As well as the incident type and the target, the other significant characteristic that must be considered in order to select a response is the likely *result(s)* of an intrusion. However, this aspect cannot be represented in only one dimension, and the taxonomy presented here considers it to be comprised of *urgency*, *severity*, *impact(s)* and *potential incidents* arising from an incident.

The *Urgency* relates to the need for timely response, and partially reflects the speed of the attack. Since some attacks can evolve more rapidly than others, it is important to consider how much time is available to respond in each case. A Denial of Service attack, launched with the use of automated scripts is an example of a rapidly evolving attack, while sniffing traffic in a Local Area Network (LAN) allows a greater window of opportunity for response, as it is likely to evolve in a longer period of time. Another dimension of the result is the *Severity* of the intrusion, which relates to the magnitude or extent of the attack. The more severe an intrusion is, the sooner it needs to be contained, in order to eliminate its impacts and the threat introduced in the system. In the taxonomy, both urgency and severity are rated on a scale of Low, Medium, High for each incident / target combination.

Apart from the urgency and severity, another aspect of the result is the consideration of the *Impact(s)* of an intrusion upon a system. The *Impact(s)* relate(s) to the asset(s) of

the system that have been compromised by the intrusion and may be observed and measured in relation to the *Confidentiality, Integrity* and / or the *Availability* of systems and data. Although in scenarios such as conventional risk analysis (Davey 1991) it is normal to rate these impacts on a sliding scale to indicate their severity, the taxonomy in the table that follows simply indicates whether there is a potential impact or not, as assignment of values would be too subjective.

The final element of the result relates to whether any further incidents are likely to be facilitated as a consequence of the initial attack. This is expressed in the taxonomy as *Potential Incidents*. For example, when sniffer software is used to capture network traffic, it is likely that the information obtained (e.g. user names and passwords) will enable attackers to log in as legitimate users at a later date and thus succeed in the masquerade. In other words, the potential incidents indicate the threat that has been introduced in the system after the occurrence of the original incident.

Having introduced the top-level elements of the taxonomy, the focus will now move to the incident categories identified earlier, as well as justifications to accompany the various ratings included in Table 3.

Information Gathering

The main characteristic of there intrusions is that they aim to collect information about a target and identify exploitable vulnerabilities. Although information gathering does not have significant impact upon a system, it carries the danger of the knowledge gained subsequently being used for launching other attacks with higher severity. Probe, Scan and Sniff are intrusions that fall into that category and will be described below.

Probe / Scan
Probe is used to access a target in order to determine its characteristics. Scan, on the contrary is used to access a set of targets in order to determine which of them have a specific characteristic. The characteristics in question aim to identify the architecture of targeted systems and networks, and usually relate to network configuration, as well as specific versions of services, operating systems and other types of software. The information obtained can subsequently enable the occurrence of incidents, such as spoofing, exploiting vulnerabilities and thus bypassing authentication, compromising software and introducing malware. The impacts relate to breach of confidentiality, as information is obtained without authorisation. Probing and especially scanning can also degrade availability, by producing large amounts of traffic when probing / scanning multiple targets. External servers as well as network components can be affected in this manner, as in both cases availability is highly important and it is those targets that are more likely to deal with that traffic.

The severity of scans / probes varies, depending on which target it is directed to. In the case of external servers and network components, which are genuinely subjected to unknown and thus untrustworthy users, they should be designed to be more tolerant with attacks of this nature. After all, within their normal activity they often provide the same nature of information anyway. Thus the severity of probing / scanning is not significant in those two cases. The urgency to respond is equally low, as apart from having low severity, probing / scanning is not likely to escalate rapidly. On the contrary, probing or scanning an internal server is not usual and thus it raises higher level of suspicion. Bearing in mind the importance of preserving confidentiality in internal servers, the level of high severity is more appropriate. The urgency to respond is medium, due to the high level of severity on one hand and its slow nature, in terms of escalating on the other. As for user workstations, although probing / scanning a user workstation is even more rare and thus raises higher level of suspicion, its impact is not as severe, as the threat to confidentiality in this case is significantly lower. Thus the severity can be regarded as ‘medium’. However, the occurrence of such an incident could mean prior breach of another target (e.g. DNS server), and thus a medium level of urgency to respond is considered appropriate.

Sniff
Sniffing consists of the interception of traffic while it travels across the network. It is achieved with the use of software tools that can capture network packets either locally or remotely. The sort of information obtained with sniffing could be anything that travels across the network, such as user name and password combinations, data files, and system or network information. After obtaining information with sniffers, the potential incidents likely to follow can mainly be masquerading, bypassing, and software compromise.

The impacts of sniffing mainly involve loss of confidentiality, however its severity and urgency depend on the type of targets subjected to it. In external servers the severity is low, since again the nature of information disclosed cannot be significant enough to raise the level of severity. Similarly with probing / scanning, the need for timely response is low, since the severity of the incident and the chance of escalating are low. In the case of internal servers, the severity is again high, however the need to respond is high as well, since the nature of information that can be disclosed in this case is more significant and thus requires a more urgent issue of response. As for user workstations, the nature of information exposed is not significant enough to increase the level of severity and urgency, so as in the case of probing / scanning, both are considered as medium.

INCIDENT	TARGET	RESULT					
		URGENCY	SEVERITY	IMPACT			POTENTIAL INCIDENTS
				C	I	A	
1.Information gathering							
Probe / Scan	External server	Low	Low	✓		✓	Spoof, Bypass, S/w compromise, Malware
	Internal server	Medium	High	✓			
	User workstation	Medium	Medium	✓			
	Net. component	Low	Low	✓		✓	
Sniff	External server	Low	Low	✓			Masquerade, Bypass, S/w compromise
	Internal server	High	High	✓			
	User workstation	Medium	Medium	✓			
	Net.component	Medium	Medium	✓			
2. Authentication failure							
Masquerade / Spoof	External server	High	High	✓		✓	Misuse, Malware, Software compromise
	Internal server	High	High	✓			
	User workstation	Medium	Medium	✓			
	Net. component	High	High	✓		✓	
Bypass	External server	High	Medium	✓			Misuse, Malware
	Internal server	High	High	✓			
	User workstation	High	Medium	✓			
	Net. component	High	Medium	✓			
3. Software Compromise							
Buffer Overflow	External server	High	High		✓	✓	Bypass, DoS, Misuse, Malware
	Internal server	High	High		✓	✓	
	User workstation	High	Medium		✓	✓	
	Net. component	High	Medium		✓	✓	
Flood / DoS	External server	High	High			✓	Spoof
	Internal server	High	High			✓	
	User workstation	Medium	Medium			✓	
	Net. component	High	High			✓	
4. Malware							
Trojan Horse	External server	High	High	✓	✓	✓	Bypass, Misuse, Malware, S/w compr., Info. gathering
	Internal server	High	High	✓	✓	✓	
	User workstation	High	High	✓	✓	✓	
	Net. component	High	High	✓	✓	✓	
Virus / Worm	External server	High	High	✓	✓	✓	Misuse, Malware, S/w compr., Info. gathering
	Internal server	High	High	✓	✓	✓	
	User workstation	High	High	✓	✓	✓	
	Net. component	High	High	✓	✓	✓	
5. Misuse							
Unauthorised Alteration	External server	High	High		✓	✓	Malware
	Internal server	High	High		✓	✓	
	User workstation	High	Medium		✓	✓	
	Net. component	High	High		✓	✓	
Unauthorised Access	External server	High	Low	✓			Malware, Unauthorised Alteration
	Internal server	High	High	✓			
	User workstation	High	Medium	✓			
	Net. component	High	Low	✓			

Table 3: Response–oriented Intrusion Taxonomy

Finally, in the case of network components, the severity of sniffing is medium, since the nature of information exposed in this case (e.g. Access Control Lists, administrator user account details) is significant enough to raise the level of severity. The urgency to respond is also medium, since network components represent single points of failure and a possible compromise could affect multiple hosts.

Authentication failure

Users and processes need to identify and authenticate themselves quite often in order to obtain specific access privileges. As a result, defeating the authentication process is very common objective for attackers, and can be summarised in three main ways, namely Masquerading, Spoofing and Bypassing.

Masquerade / Spoof
Masquerade is the action in which valid identification and verification information that belongs to legitimate users is obtained and used by an impostor. For example, an attacker might use a sniffer to capture user name, password and IP address combinations that are sent across the network, and then use this information to log into accounts that belong to other users. Spoofing, by contrast, involves the provision of false information. In network communications, each packet of information traveling on a network contains source and destination addresses either in the form of MAC, IP addresses, TCP connection IDs, or port numbers. Supplying accurate information is often assumed, however it is possible that incorrect information is entered into these communications, in order to accept an impostor address as original and either trick other machines into sending it data or to allow it to receive and alter data. Examples include IP spoofing, email spoofing and DNS spoofing.

Masquerading and spoofing are mainly a threat to the confidentiality of systems, since they most often provide unauthorised increased access to attackers. However, in the case of external servers and network components, it is possible to cause loss of availability as well, if used as a technique to enable the occurrence of DoS attacks. The potential incidents that can follow masquerading and spoofing are obviously misuse (unauthorised access and alteration of information), malware (introduction of Trojan horses, viruses / worms) and software compromise (Buffer overflow, DoS).

The severity of masquerading and spoofing is considered high in external servers, as it may result in loss of availability. The urgency to respond is high as well, since IP spoofing can very soon escalate to a DoS incident. However, even in the case of masquerading, once unauthorised access is achieved to external servers, it is possible to alter information that can harm the public image of the organisation and thus cause further embarrassment and disruption of operation. In the case of internal servers, even if services are not accessed externally, the danger of disclosing confidential information is considerably high, resulting in severe embarrassment to the organisation, and disruption of its operation. So, the level of severity and the urgency to respond in this case are high as well. As for user workstations, the severity is less significant, as in many cases the nature of information or access level obtained will not pose a great level of threat to the system (although some users will always be exceptions). The level of urgency is medium as well, since the workstation is probably used as a step to achieve increased access into a more significant component of the system (either internal or external server). Obtaining unauthorised access in network components, as well as making them unavailable by achieving DoS attacks is highly severe, as it can affect multiple hosts or even the entire internal network, depending on the scale of the problem. The urgency to respond is thus high as well.

Bypass
Bypass is an action taken to avoid the authentication process by using an alternative method to access a target. For example, some operating systems have vulnerabilities that could be exploited by an attacker to gain privileges without actually logging into any privileged account. Bypass is usually a result of software compromise (e.g. buffer overlow) or malware (e.g. if a trojan horse is used instead of the original authentication process). The issue is again a threat to confidentiality, as increased unauthorised access is achieved. The potential incidents that can follow are misuse (unauthorised access and alteration of information) and malware.

The severity is medium in the case of external servers, since their availability is not threatened directly. However a rapid response is needed to avoid further escalation of the incident, so the urgency in that case is high. In internal servers both severity and urgency are high, as the direct threat is higher, so is the need to avoid escalation of the incident. Although the severity in the case of user workstations is lower, and thus can be considered as medium, the need to respond is equally high, since bypassing authentication is an indication of an already compromised system, so further action should be taken as soon as possible. Finally, bypassing authentication in network components is of medium severity, since the threat to confidentiality is not as severe as in the case of internal servers, but again the need to respond and eliminate any chances of escalating the problem is high.

Software compromise

Intrusions that involve the exploitation of software vulnerabilities fall into this category. There are three

main categories of vulnerabilities within a system, namely design, implementation or configuration vulnerabilities (Howard 1997). The main categories of intrusions that fall into this category are Buffer Overflow and Denial of Service; they are presented below.

Buffer Overflow
Buffer overflow is a result of deficient software implementation that allows the assignment of data in a buffer without checking in advance if its size is sufficient to 'host' that data. So in the case of someone sending larger amounts of data, the targeted system will allow the input of data in the buffer anyway, with the result of either crashing the system or overwriting part of memory adjacent to the buffer. As a result of the latter, unauthorised access could be obtained by modifying the flow of program execution, and allowing the execution of arbitrary code with the same access rights granted to the compromised program (Aleph1 1996).

Such incidents can compromise the integrity and availability of the targeted system, and can lead to further incidents such as bypassing authentication, denial of service, misuse or execution of malware. In all cases, the amount of time elapsing before that happens is usually small, as in many cases it even happens almost simultaneously.

Buffer Overflows are more commonly exploited in server software (web, ftp, email, file) since they are easily accessible from external sites and often run under root/administrator privileges. Thus high potential severity can exist for external servers, as well as internal (intranet) servers in some organisations. The urgency to respond is high as well, since apart from the significant severity of the incident, the likelihood of escalation is significant as well, so an urgent response is needed.

In the case of user workstations the severity is medium, since the chances of being subjected to attacks of this nature is less substantial. Also, even if targeted (e.g. server software is running, probably by default) the number of hosts affected are limited (probably only one), so the scale of the problem is less significant. However, the urgency to respond is still high, in order to avoid execution of malware or further compromise of other systems.

The chance of exploiting buffer overflows in network components is even less significant, but the potential impacts of doing so are more substantial than in the case of workstations, since a greater number of hosts can be affected. Thus the severity of buffer overflow is medium in this case. The urgency to respond is again high, for the same reason.

Flood / Denial of Service
Denial of Service (DoS) attacks aim to overload (flood) the capacity of a target by accessing it repeatedly. The result of such action is to make the target unable to respond to any other events / requests and thus become inaccessible to legitimate clients. Subsequent occurrences could include another party assuming the role of the target, resulting in spoofing.

The impact of Denial of Service attacks clearly relates to the availability of the targets. Since these attacks are most often conducted with the use of automated scripts, the need to respond immediately is crucial in most cases. In the case of an external server, the severity is likely to be high, given that a site may represent a public interface of the organization. Inaccessibility could result in embarrassment and loss of custom. The urgency to respond is also high, since usually the time available to prevent either the occurrence of the incident, or subsequent escalation, is very limited. Although DoS to internal servers and network components does not risk causing embarrassment to the organisation, their failure to provide services could have impact on multiple hosts, or even the entire internal network of the organisation, so the severity is also high, as is the urgency to respond. In the case of user workstations, the likelihood of being subjected to a DoS attack is rather small, simply because the impact of doing so is not as significant. User workstations are mostly used as (potentially unwitting) tools to conduct DoS attacks in order to achieve maximum level of effectiveness, but are not the targets. However, it is possible, and it can result in either degradation of performance, or total loss of legitimate usability. Thus the severity in that case is medium. The urgency to respond is medium as well, as the impacts of the attack are of medium severity and the time available to encounter the attack or avoid escalation is usually more.

Malware

Malicious software, also known as malware, characterises the classes of intrusions that are conducted under complete software control. Intrusions falling into this category differentiate from automated software tools used to launch other classes of attacks (e.g. DoS attacks), in the sense that humans are not involved in the escalation of malware attacks; after the initial human involvement to begin the distribution of malware, individual attacks can subsequently occur without the need for the instigator's further involvement. Thus malware can constitute an attack in its own right. There are three main types of malware, namely Trojan horses, viruses and worms and will be discussed below.

The impacts of malware can differ significantly from case to case, since the code in the payload can do nearly

everything that is feasible under software control. For example, it is possible to initiate posting of legitimate users' working documents to all the members of his/her address book, resulting to breach of confidentiality (SARC 1999). Alternatively, it is possible to delete or modify files in the system, achieving a breach of integrity. Finally system or network resources can be consumed at the execution of the payload, resulting to either degradation of performance or entire inaccessibility of targets for legitimate use.

The potential incidents that can follow the execution of malware can also be nearly anything. Misuse, other forms of malware, software compromise and information gathering are examples of potential results of malware. Thus the severity of malware varies according to the specific incidents. However, if considering the execution of malware in general, the severity is high in all types of targets, since such a great variety of functionality can potentially be included in the payload. In addition, the risk of spreading to additional targets is extremely high, so the urgency to respond and contain the execution of malware is high as well in all cases.

Misuse

Misuse relates to unauthorised or unacceptable use of system resources. In this sense, it is a quite general term that can actually include all the incidents described so far, since all of them are somehow a form of misusing system resources. However, incidents falling into this category mainly take place after unauthorised access has been obtained in a target and include cases that mainly involve misuse of files and data within a system. It is important to mention at this point that the occurrence of incidents from this category indicates that the targeted system may have already been in a compromised state, unless the activity is being perpetrated by a legitimate user. Hence any response issued might be affected by this factor as well.

Unauthorised alteration
Unauthorised alteration includes actions such as creating, modifying, deleting system or data files. This will affect the integrity and / or availability of resources and represents an important issue that needs to be addressed.

The severity in the case of external servers is high, as information or services might be altered in such a way as to cause embarrassment to an organisation and further disruption to its normal operation. For example, web site defacements (Alldas.de 2001) represent a highly important incident that can immediately attract the interest of media and put the organisation into a difficult situation. Also the modification of information or services could potentially mislead or cheat customers, and result in making the organisation liable for those actions. Although the urgency to respond in such case is high, the feasibility of doing so might be another issue. Certainly the current state of the system needs to be considered in order to determine the effectiveness or selection of an appropriate response.

Unauthorised alteration is highly severe in the case of internal servers and network components as well, since it can result in misleading internal users to make decisions based on inaccurate information or disrupting their operation. Even if the likelihood for rapid escalation of the incident is very small, the need for timely response is high again, since the severity of the incident is so significant.

Finally in the case of user workstations, the importance of the target is typically lower, as it can affect only a limited number of users. The severity is therefore medium. Still, the urgency to respond is high, mainly because the current state of the targeted system should be assessed and any potential risks minimised.

Unauthorised Access
Unauthorised access includes actions that involve disclosure of information to unauthorised parties. As a result of their occurrence, incidents such as unauthorised alteration or execution of malware might follow. Thus the severity of unauthorised access can vary according to the target (and whether confidentiality is at high risk) but the urgency to respond in all cases should be high. That is to firstly assess the current state of the system and prevent further escalation of the incident and occurrence of unauthorised alteration or execution of malware as well.

When external servers or network components are subjected to unauthorised access, the severity is low, since no confidential information should be at risk and no modification has taken place. On the other hand the current state of the system is unknown and needs to be assessed. By contrast, unauthorised access to internal servers has high severity, because there is more important information available for attackers. In the case of user workstations the severity is medium, as there is risk to confidentiality, but it is less substantial.

CONCLUSIONS

In this taxonomy, several incidents have been considered, aiming to illustrate the effect of different types of targets on the results of an intrusion. The ultimate intention is to give insight into the main intrusion characteristics that can influence intrusion response, and subsequently lead to the indication of generic classes of response. Although the response-oriented taxonomy is quite generic and cannot depict the complexity of the response decision process, it can still serve as a basic tool that will enable the research to progress towards that direction. After looking into the results of different intrusions on various targets, it seems

that intrusions directed towards internal servers always have the most significant results, mainly due to their importance in the operation of an organisation. By contrast, user workstations have the least significant results, as their role within the organisation is less important and the consequences after the occurrence of an intrusion can more easily be addressed. Finally, network components and external servers seem to depend on the type of intrusion to a greater extent, as some classes of intrusions have more significant results than others.

In terms of response and how different intrusion characteristics can influence the response process, it can be argued that the more severe an intrusion is, the more important it is for the response to focus on the prevention of its occurrence, or its containment. In classes of intrusions with low or medium severity and high urgency, the risk for rapid escalation is significant, and so the response process should focus on the prevention of further escalation (prevent the occurrence of potential incidents). Finally, the severity and urgency can affect the transparency of the initiated response. It seems that there should be a trade-off between them, as the more severe the intrusions, the less transparent responses can apply.

It should be noted that there are several limitations in this taxonomy. For example, apart from the type of target, the number of systems targeted could also be considered, as the scale of an incident will certainly influence its severity. For example, a virus that infects a small number of user workstations is not as severe as one that infects all of them. However, the omission of this factor does not prevent the taxonomy from fulfilling its objective of demonstrating that the same category of incident can demand different responses in different contexts.

As regards the responses themselves, it may appear curious that they have been omitted from the taxonomy presented here. The basic reason is that the taxonomy is intended to provide the foundation for an automated decision mechanism within a software agent. The specific response options available could vary depending upon the environment in which the agent is deployed, and thus the classification taxonomy is independent of any particular mapping. In the context of such an agent, the decision-making process could also be more complex. Although incident and target related characteristics are the main determinant of the likely result of the incident, various other contextual factors could be measured when an incident is detected in order to better inform the response decision process. For example, the account in use, the current alert level of the IDS, and the nature of any responses already issued could all influence the choice of response that is likely to be the most effective. Further consideration of this issue is presented in (Papadaki et al. 2002), and the issue represents the focus of ongoing research by the authors.

REFERENCES

Aleph1 (1996), “Smashing The Stack For Fun And Profit”, Phrack online journal, vol. 7, issue 49, 8 November 1996.

Alldas.de (2001), “Defacement Archive”, http://defaced.alldas.de/, 26 October 2001.

Cheswick W.R., and Bellovin S.M. (1994), 'Firewalls and Internet Security: Repelling the Wily Hacker', Addison-Wesley Publishing Company, 1994.

Cohen F.B. (1995), *Protection and Security on the Information Superhighway*, John Wiley & Sons.

Davey J. (1991), “The CCTA risk analysis and management methodology (CRAMM)”, Current Perspectives in Healthcare Computing, pp. 360 – 365.

Furnell S.M. and Dowland P.S. (2000), "A conceptual architecture for real-time intrusion monitoring", *Information Management & Computer Security*, Vol. 8, No. 2, pp65-74.

Furnell, S.M, Magklaras, G.B, Papadaki, M. and Dowland, P.S. (2001), "A Generic Taxonomy for Intrusion Specification and Response", in *Proceedings of Euromedia 2001*, Valencia, Spain, 18-20 April 2001: 125-131

Howard J.D. (1997), PhD thesis 'An Analysis of Security Incidents on the Internet 1989 - 1995', Carnegie Melon University, 7 April 1997, http://www.cert.org/nav/reports.html

Lindqvist U., and Jonsson E. (1997), “How to Systematically Classify Computer Security Intrusions”, in Proceedings of the 1997 IEEE Symposium on Security and Privacy, May 4-7, 1997, IEEE Computer Society Press.

Neumann P.G., and Parker D.B. (1989), “A summary of computer misuse techniques”, in Proceedings of the 12th National Computer Security Conference, Balitimore, USA, 10-13 Oct 1989, pp. 396-407.

Papadaki, M., Furnell, S.M., Lee, S.J., Lines, B.M. and Reynolds, P.L. (2002), “Enhancing response in intrusion detection systems”, submitted to *Journal of Information Warfare*.

Russell D. and Gangemi G. T. (1991), "Computer Security Basics", O’Reilly & Associates, Inc., Sebastopol, CA, 1991.

SARC. 1999. “W97.Melissa.A virus overview”. Symantec AntiVirus Research Center. http://service1.symantec.com/sarc/sarc.nsf/html/W97.Melissa.A.htm

ACKNOWLEDGEMENTS

The research presented in this paper has been supported by funding from the State Scholarships Foundation (SSF) of Greece.