

## Automated Intrusion Response

a report by

**Dr Steven Furnell** and **Maria Papadaki**

*Network Research Group, School of Computing, Communications & Electronics, University of Plymouth*

### Introduction

The increasing speed of attacks against information technology (IT) systems highlights a requirement for comparably timely response. Threats such as malware and scripted exploits often allow a timeframe of only a few minutes or even seconds to respond, which effectively eliminates the feasibility of manual intervention and highlights a requirement for automated approaches to provide a solution. Here, however, it can be seen that existing security technologies are often insufficient. For example, although intrusion detection systems (IDS) can be used to identify potential incidents, they have a tendency to produce high volumes of false alarms and consequently cannot be trusted to issue automated responses for fear of disrupting legitimate activity. The inability of IDS to directly tackle intrusions was the main criticism of a market report released by Gartner in June 2003, which labelled them a “market failure” and predicted their obsolescence by 2005. Indeed, more recent years have witnessed an apparent shift in the popularity of the associated security products, with intrusion prevention systems (IPS) gaining increasing appeal over IDS solutions.<sup>2</sup> IPS products themselves can actually use similar underlying detection methods to the IDS approach, but differ in that they attempt to sidestep the problem of false-positives by responding only to attacks that can be detected with high certainty. The confidence that results from this allows the IPS to be placed in-line, between the source of an attack and the potential victim, giving it the potential to directly prevent incidents by blocking offending traffic. However, even if this is an important step in the right direction, it still leaves the problem of all the other attacks, which are more difficult to detect, and which are allowed to pass with no response. When viewed in this sense, it becomes apparent that IPS should not be seen as an alternative to IDS, but as another layer of security within a defence-in-depth strategy. As such,

this brings back the problem of trusting an IDS to make reliable response decisions.

### Enhancing and Automating the Intrusion Response Process

Existing IDS do, of course, have some level of automated response capability. The constraint, however, is that these are often limited to passive actions, such as logging and raising alerts to request manual investigation, which do nothing to directly impede the progress of an intruder. To achieve the latter, active responses have been introduced that mainly focus on blocking or terminating offending events. Such response actions might include:

- limiting or denying user access;
- blocking network traffic through firewalls and routers; and
- terminating network connections.

However, in face of the false alarms, security administrators are naturally wary of automating such responses in current IDS. The following comment was typical of those received from an opinion survey conducted among relevant security professionals and related product vendors:

*“Proactive measures are a reasonable idea, unless they can be subverted. For instance, if you decide to shut down your network connections as a proactive approach, then an intrusion attempt can be used as a denial of service.”<sup>3</sup>*

As a consequence, it is useful to consider approaches that may be used to mitigate the risk of issuing active responses in false alarm scenarios. The key requirement here is to take into account the possibility of having a false alarm, and attempt to reduce the adverse impacts of automated response. One approach

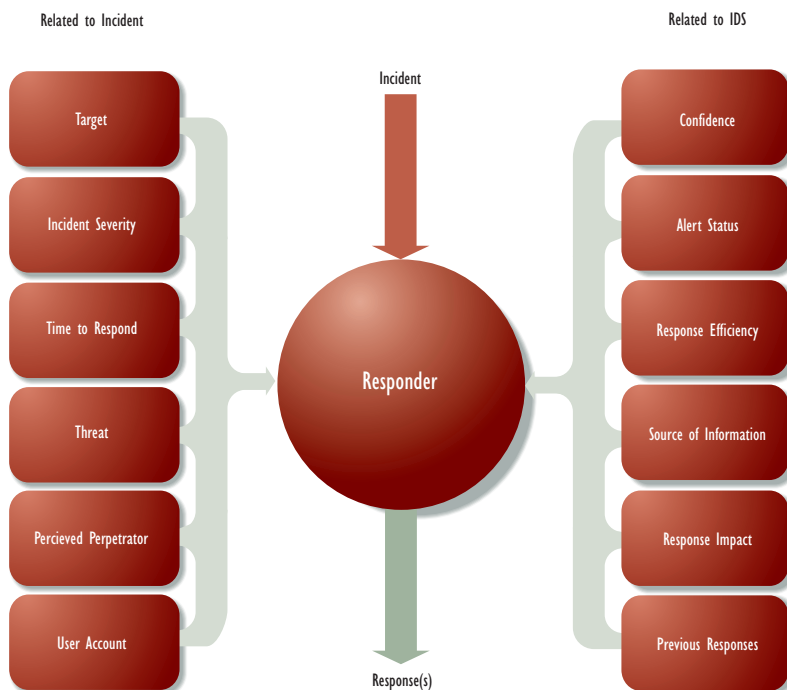
Dr Steven Furnell is Head of the Network Research Group at the University of Plymouth, UK. He has been actively involved in security research for over 13 years, and has authored numerous papers on the topic, as well as the books *Cybercrime: Vandalizing the Information Society* (Addison Wesley) and *Computer Insecurity: Risking the System* (Springer).

Dr Maria Papadaki is working as a security analyst for Symantec in the UK. She recently completed her PhD research within the Network Research Group, focusing on the issue of flexible, automated intrusion detection system (IDS) response. This research activity was undertaken with support from the State Scholarship Foundation of Greece.

1. Gartner, “Gartner Information Security Hype Cycle Declares Intrusion Detections Systems a Market Failure”, Gartner Press Release 11 June 2003.
2. Gordon L A, Loeb M P, Lucyshyn W, Richardson R, “Ninth Annual CSI/FBI Computer Crime and Security Survey”, Computer Security Institute, (2004).
3. Papadaki M, Furnell S M, Lee S J, Lines B M, Reynolds P L, “Enhancing response in intrusion detection systems”, J. Information Warfare (2003), 2; 1: pp. 90–102.



Figure 1: Contextual Factors Influencing Intrusion Response



is to concentrate on informing decisions as much as possible, and enabling a responder to issue actions that investigate attacks, collect more evidence, or postpone or delay the attack while investigating, limiting the effects of an attack or a response at the target. So, in addition to automation, two further desirable characteristics can be identified:

- Flexibility – The recognition that the type(s) of response that are appropriate will often depend upon the context in which an incident has occurred, such that the same incident will often demand a different response.
- Intelligence – The capability to assess the appropriateness of response actions before and after initiating them.

Accepting these requirements leads to the natural question of how they can be achieved, and it is here that the discussion enters the domain of on-going research. However, given that the detection methods in many of today's commercial IDS approaches have their origins in research activities dating back to the 1980s, it is relevant to consider how current research may help to advance commercial incarnations in the future.

### Establishing the Context of an Incident

In order to choose the most appropriate response action(s), it is relevant to consider the context in which an incident is occurring. For example, a malware infection would require different responses on a standard end-user workstation than it would on

a critical database server. In the first case, the priority would be to keep the malicious code from spreading, and hence a suitable response would be to disconnect the host from the network. In the second scenario, however, maintaining the operation of the server would be important, so other response actions might be selected, such as quarantining the malicious code, backing up the database, and temporarily restricting the execution of any processes on the system that are not relevant to database transactions. In this example, the only factor being assessed is the target of the attack. However, it is possible to identify a whole range of factors that may influence the response decision. Some indicative examples are illustrated in *Figure 1*, grouped according to whether they are related to the incident or the IDS.

Although the occurrence of an incident obviously remains the trigger for a response, and still represents the principal influence over what should be done, the assessment of other factors enables the responder to establish the context in which the incident has occurred, and therefore select appropriate responses accordingly. The various factors related to the incident are defined in more detail below.

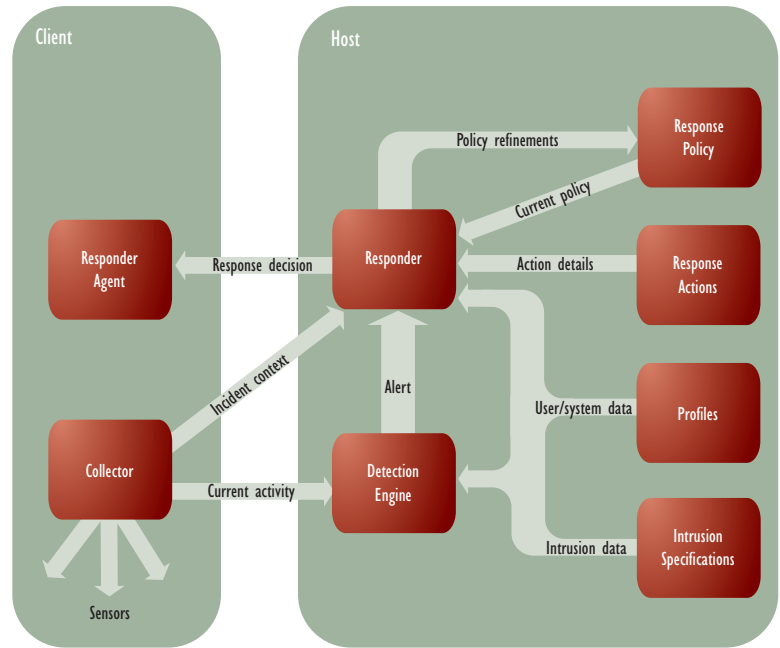
- Target – What system, resource or data appears to be the focus of the attack? What assets are at risk if the incident continues or is able to be repeated? How important is that resource for the continuation of the system's operation?
- Incident severity – What impact has the incident already had on the confidentiality, integrity, or availability of the system and its data? How strong a response is required at this stage? For example, the detection of a severe incident could warrant the initiation of correspondingly severe responses, in order to protect system resources.
- Time to respond – How urgently is a response needed? This factor is mainly influenced by the speed of the attack.
- Threat posed by incident – How serious is the threat to the system after the occurrence of the incident? Which attacks are more likely to follow after that incident?
- Perceived perpetrator – Does the evidence collected suggest that the perpetrator is an external party or an insider? Is there any history associated with that person/account?
- User account – If the attack is being conducted through the suspected compromise of a user account, what privileges are associated with that account? What risk do those privileges pose to the system?

Factors related to the IDS are summarised below.

- Confidence – How many monitored characteristics within the system are suggestive of an intrusion having occurred?
- Alert status – What is the current status of the IDS, both on the suspect account/process and in the system overall?
- Response efficiency – What has the efficiency of a specific response proven to be under specific conditions? The efficiency rating of a specific response can be updated after considering its efficiency in previous instances of the same incident.
- Source of information – What is the detecting capability of the source of information about the incident? Some sources or IDS metrics might be more reliable in detecting attacks than others, generating fewer false positive alarms (e.g. anomaly detectors tend to generate more false positive alarms than misuse detectors,<sup>4</sup> and some monitoring sensors produce fewer false alarms than others, depending on their location and configuration). The responder should be able to determine the credibility of sources over time and use this to inform response selection.
- Response impact – What would be the impact of initiating a particular form of response? How would it affect a legitimate user if the suspected intrusion were, in fact, a false alarm? Would there be any adverse impact on other system users if a particular response was taken?
- Previous responses – If one or more responses have already been issued as a result of the detected incident, and been unsuccessful in countering the intrusion, it would be relevant to consider this before determining the acceptable impact of the next action. The failure of previously issued responses might lead to the selection of more severe response actions.

One advantage of considering such factors is the aforementioned flexibility. In addition, establishing the incident context is also recognised as a means of limiting uncertainty in the response process,<sup>5</sup> and can be a contributor to reducing the potential impacts arising from IDS false alarms.

Figure 2: The FAIR Architecture



#### A Framework for Automated Response

Having identified the requirements, it is necessary to consider a means by which they can be achieved. The authors' research has proposed the conceptual framework for a flexible automated intelligent responder (FAIR),<sup>6</sup> based on the concept of a co-ordinating host handling the monitoring of a number of networked client systems. The core elements are illustrated in *Figure 2*, and summarised below.

- Detection engine – analyses current activity and raises alerts for suspected intrusions. Informs the responder of the intrusion type, along with factors such as the target of the attack, and perceived perpetrator.
- Responder – monitors alerts, considering them in conjunction with incident context to take appropriate actions where necessary.
- Collector – provides initial activity data to detection engine, and subsequently informs responder about current context on the target system (e.g. applications running, active network connections, processor load).
- Intrusion specifications – intrusion specifications contain information about specific types of

<sup>4</sup> Bace R, Mell P, 2001 NIST Special Publication on Intrusion Detection Systems, National Institute of Standards and Technology (NIST), <http://csrc.nist.gov/publications/nistpubs/800-31/sp800-31.pdf>

<sup>5</sup> Carver C A, Jr, Hill J M D, Pooch U W, "Limiting Uncertainty in Intrusion Response", 2nd Annual IEEE Systems, Man, and Cybernetics Information Assurance and Security Workshop, West Point New York, June 5–6 2001.

<sup>6</sup> Papadaki M, "Classifying and Responding to Network Intrusions", PhD Thesis, (2004), University of Plymouth, United Kingdom.

intrusions and their characteristics, such as incident severity rating, estimates of likely impacts (e.g. in terms of confidentiality, integrity and availability), and the speed with which the attack is likely to evolve.

- Profiles – contain data about users, systems and attackers, which can provide additional context for response decisions.
- Response actions – details of available response actions, enabling selection of responses with the most appropriate characteristics (e.g. stopping power, intrusiveness).
- Response policy – uses expert systems technology to indicate the most desirable characteristics for responses in the current context.
- Responder agent – initiates and manages any response actions required on the target (e.g. correcting vulnerabilities, authentication challenges, limiting access rights).

As seen in *Figure 2*, the responder uses information from several sources to determine the context of an attack, which informs the ultimate response decisions.

In terms of bringing such research forward to operational reality, elements of the FAIR architecture have been realised within an initial proof-of-concept prototype. This demonstrates many of the core elements of the architecture and its associated assessment of contextual factors leading to the ability to show flexible response decisions being made on the basis of simulated attacks. However, further work is required to incorporate the adaptive learning capability that would enable automated refinements to the response policy based on experience over time. In addition, a comprehensive operational evaluation of the concept would still be required in order to demonstrate the reliability and impact of the system. Nonetheless, the pursuit of such research is clearly warranted by the threats that are posed to systems and the risks of not responding to them in an appropriate and timely manner. ■