AR-010-350

DSTO

Directions for Intrusion detection and Response: A Survey

Dean Engelhardt

DSTO-GD-0155

APPROVED FOR PUBLIC RELEASE

DEPARTMENT OF DEFENCE

DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION

DIRECTIONS FOR INTRUSION DETECTION AND RESPONSE: A SURVEY

Dean Engelhardt

Information Technology Division
Electronics and Surveillance Research Laboratory

DSTO–GD–0155

ABSTRACT

This document presents a review of recent research into the automated detection of attack on computers or networks of computers. It is now widely regarded that despite efforts to secure computer systems against intruders by operating system protection, an increasing number of such attempts are succeeding. Only through the careful monitoring of activity in a computation environment can such penetrations be detected and potentially repelled. We briefly describe some early attempts to provide such monitoring and then proceed to describe several research efforts that are currently underway to overcome the limitations of these classical systems. From an analysis of these new approaches we distill several core principles that are critical to the success of future detect and react systems.

APPROVED FOR PUBLIC RELEASE

DEPARTMENT OF DEFENCE

DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION

Published by

DSTO Electronics and Surveillance Research Laboratory
PO Box 1500
Salisbury, South Australia, Australia 5108

Telephone: (08) 8259 5555
Facsimile: (08) 8259 6567

AR No. AR-010-350
October, 1997

APPROVED FOR PUBLIC RELEASE

Directions for Intrusion Detection and Response: A Survey

EXECUTIVE SUMMARY

In recent years it has become increasingly apparent that attacks on the computing facilities of an organisation pose a potentially serious threat. Numerous stories have emerged describing security breaches occasioned by so-called hacker attacks in which individuals electronically 'break-in' using a computer situated in a geographically remote location. Such attacks have the potential to expose sensitive information to unauthorized eyes, or even worse permit unauthorized modification or sabotage of a computer and its stored information. Evidence suggests that, despite efforts to provide better protection to our computers (usually in the form of better operating system security mechanisms), the number of penetration attempts — and the number of successes — is increasing. This trend has very serious ramifications to both the civilian and defence communities, since the modern operation of both are increasingly reliant upon the correct operation of computer-based systems.

One approach that has been adopted in an effort to overcome some of these serious problems, is the construction of automated software tools for detecting an intrusion (preferably as it is occurring) and responding to it. A number of simple systems which implement such functionality exist already, although considerable research is ongoing to improve the quality of automated detection and the ability to effectively respond. This document surveys several such research efforts.

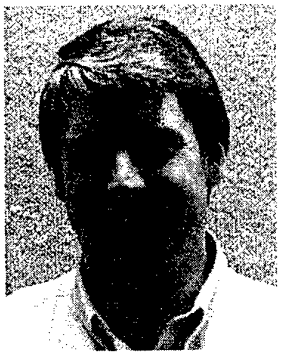
The proposals and prototypes described herein can be broadly categorised as attempts to address three key areas in which the state of the art is deficient, namely:

- in providing protection for very large networks of computers,
- in reliably tracing an intrusion attempt back to its origin, and
- in providing artificial intelligence techniques for reliably detecting intrusions.

From an analysis of the techniques employed in the various research projects we consider, coupled with a consideration of the motivations behind them, several properties can be distilled as highly desirable elements of future detect and react systems. These include:

- the capacity to scale to protect computer networks of any size,
- an effective capacity to respond to an intrusion with a high degree of confidence that the response is not misdirected, and
- a graceful degradation in the face of failure or compromise.

Authors



Dean Engelhardt
Information Technology Division

Dean Engelhardt joined the Advanced Computer Capabilities Group (ACC) as a Research Scientist in April of 1997. Prior to this he worked and studied at the University of Adelaide, completing a Ph.D. in the Faculty of Engineering.

Contents

1 Introduction

This document reviews recent research trends in the field of computer intrusion detection and automated response. We briefly describe classical approaches to providing security based upon these types of functionality, noting the deficiencies in the current state of the art. Throughout the remainder of this paper we describe and analyze a number of attempts to overcome these problems by the application of new techniques.

We begin by noting the research priorities for detect and react which have been laid down in recommendations to funding bodies, describing the particular features that current INFOSEC agencies believe are critical to the utility of future systems. We then survey several current research projects which seek to better the state of the art in intrusion detection and automated response. Our analysis is broken down by geography, with the North American survey being further divided into those projects which are being funded by the Defence Advanced Research Projects Agency (DARPA) and those which have no direct defence affiliations. We conclude by drawing common threads from the systems we have reviewed, isolating the aspects which seem most promising for incorporation into future research.

2 Classical Intrusion Detection

Despite the relative newness of intrusion detection as an area of study (due to the fact that it is only recently that potential threat posed by hackers and viruses has become widely recognized), there already exist several moderately mature systems which attempt such functionality. By and large these environments, some of which are touted as commercial products, are derived from the model of intrusion detection first proposed by Denning [4], and used by Lunt *et al.* in the construction of the Intrusion Detection Expert System (IDES) [12] and its next-generation successor NIDES [11]. This approach can be characterized by the fact that it works directly upon audit trails of user activity, attempting to locate patterns or trends which are indicative of an attempt to breach the security of the machine. Typically, two different approaches are simultaneously taken in such detection: the utilization of an expert system to match patterns in the audit trail against known attack signatures, and the profiling of user command use to detect noticeable departures from historically recorded profiles. Each method attempts to isolate a different type of security threat — the former protects against *misuse* by a hacker who is aware of holes in the system security, while the latter aims to note *anomalies* in a user's machine use which may be indicative of either a successful hacker intrusion or an insider attack. It is necessary to use both forms of detection simultaneously since neither has the ability to expose all possible threats, and each by itself can be readily tricked by a knowledgeable attacker.

Classical intrusion detection systems, as epitomized by IDES and NIDES, while they offer a basic level of ability to detect misuse and anomalies, are rudimentary tools which provide little protection against many alternate attack scenarios. This is particularly true in the case of networked computers, where a threat may span many machines in a way which is seemingly innocuous from the perspective of any one machine but which

poses serious danger to network-wide security. To detect such attacks, an intrusion detection system must be able to see a bigger picture than merely the audit details of a single machine. More recent systems (such as the Distributed Intrusion Detection System (DIDS) [18] and the Network Anomaly Detection and Intrusion Reporter (NADIR) [7]) constitute attempts to provide such a view. Typically this is achieved by maintaining some centralized resource, which collects details from the networked machines and builds up a coherent large scale picture of the potential threat. This centralized approach is inherently limited in its scalability, typically only allowing for consideration of fairly small networks. Indications are that in the future intrusion detection systems will be called upon to provide network-level protection at a very much larger scale, for example across Wide Area Networks (WANs).

Current intrusion detection systems, even those catering to network-level co-ordination, still have many flaws. Chief amongst these is their very limited capabilities of response: most limit themselves to reporting suspicion to a human operator by email or graphical user interface. Clearly in a modern environment where an administrator may only need to check upon the system once per day, this approach offers no possibility for rapid response (i.e., "closing the breach" quickly). This time lag to respond, coupled with the large bandwidth of modern networks, gives potential for vast quantities of information to leak as a result of a successful intrusion. Another problem which remains in state of the art intrusion detection systems is their reliance on simplistic forms of statistical and/or expert-system analysis, each of which is open to defeat by well-understood techniques (e.g., the slow alteration of user profile, or the minor alteration of attack from a recognized signature). Finally, as noted above, current systems cannot be usefully deployed across networks of more than a dozen or so computers, thus making them inappropriate for protecting large systems from attack.

Throughout the remainder of this report, we will concern ourselves with efforts to overcome these limitations in the current state of the art. We begin by providing an account of problems the INFOSEC community at large believe need critical attention in such systems, then progress on to describe several current efforts which seek to address these and other issues.

3 Current Funding Priorities for Intrusion and Detection Research

3.1 The National Info-Sec Technical Baseline (NITB)

As part of an effort to focus U.S. government research into technologies for information survivability, information security, and the protection of the national information infrastructure, the INFOSEC Research Council[1] seeks to establish technical baselines for several areas of INFOSEC research. The intent is that each such baseline will document

[1]This council is a newly-formed body made up of agencies such as NIST, DARPA, DISA, NSA, DOE, IDA, ONR and the CIA and several research laboratories including Naval Research Labs (NRL), Air Force Information Warfare Center (AFIWC), Rome Laboratory, Space and Naval Warfare Center (SPAWAR), and the U.S. Army Communications-Electronics Command (CECOM).

priority areas for research within an INFOSEC discipline towards which funding should be preferentially directed.

The first research area to be reviewed in this way by the INFOSEC Research Council has been that of intrusion detection and response. A draft NITB arising from this study has recently become publicly available [10]. In several ways, this document is critical of the state of the art on computer intrusion detection technology, arguing that important gaps remain that should ideally be addressed in future research. The key deficiencies cited are:

- **Lack of Basic Definitions and Mathematical Understanding:** at present there is no sound theoretical or mathematical basis for deciding what constitutes an intrusion.
- **Lack of Common Consistent Information from Multiple Audit Sources:** systems which attempt to provide intrusion detection across heterogenous networks of computers must typically synthesize a consistent picture of the dynamic state of the network from audit information which differs in format and content for each type of machine. This is cited as a major complexity in the task of effective intrusion detection, which the NITB argues may be eased by defining standards.
- **Lack of Protection Testing:** few intrusion detection systems are tested against a wide range of threats. Furthermore, many systems which are based around known attacks are surprisingly vulnerable to minor variations of the known threat. These factors can lead to systems offering a false sense of security.
- **Reduced Protection against Knowledgeable Intruders:** most of the methods surveyed by the report had at least some properties that could be exploited by intruders with knowledge of the particular technique employed, allowing them to bypass the intrusion detection system. Considering that one of the goals of such systems is to protect against insider threat, this is viewed as a deficiency.
- **No Mechanisms for Damage Assessment and Recovery:** current intrusion detection systems provide little assistance in cleaning up after an intrusion has been detected. Automated analysis tools could potentially offer considerably more assistance in this process.
- **Poor Scalability of Solutions:** to date, the field of intrusion detection has limited itself to single-computer systems or small networks. Many of the current technologies make use of centralized resources which preclude scaling to larger networks. Considering the present and future requirements of securing very large networks (e.g., an entire national infrastructure), this is viewed as a serious limitation of the technology which should be promptly addressed.

The NITB also specifically addresses technologies for automated response to detected intrusions, noting that such systems are critical to maintaining confidentiality on a computer network. The time scale at which intrusions occur makes real-time response by an operator impossible, and the massive bandwidth of modern networks means that even a small delay in securing a breach can result in the leakage of many megabytes of information.

Of the systems surveyed in the report, most are severely limited in possibilities for automated response by the nature of the audit information collected. Furthermore, a limited ability to accurately identify the source of an intrusion is a further hurdle to the integration of more aggressive styles of automated response. Three principal research areas are cited for response technology, namely:

- the limitation of the effect of automated response so as to prevent cascade and livelock failures,
- the provision of safeguards against false-positives and enemy-induced responses,
- using the response system to push the point of attack deflection back towards the attack source.

The conclusions of the NITB ultimately paint intrusion detection and response as a viable area of future research, although the recommended directions for such research have some divergence from previous approaches. Particularly advocated is deeper scientific research into some of the fundamentals of detection and reaction.

In general the recommendations embodied in the NITB are reasonably sound although lacking in insight on the specifics of what is and is not technologically viable at present. From a practical perspective, the most important points to emerge from the analysis are that the general thrust of U.S. governmental spending on intrusion detection and response will likely be directed at proposals with certain properties, namely scalability, defence against knowledgeable attack and the ability to perform damage assessment and/or recovery.

4 Current Research Projects

4.1 DARPA Directed Research

4.1.1 University of California at Davis

The largest of the DARPA/ITO funded research ventures, the *Intrusion Detection for Large Networks* project at UC Davis considers several distinct technological approaches to enforcing security and confidentiality on national-scale or global-scale networks.

Privileged Programs

One aspect of this work [9] considers the problem of detecting intrusions which are based on attempts to exploit faults or trapdoors in security critical programs (e.g., Unix programs with `setuid root`). Typically, while such programs operate in an environment where they have potential access to an extensive set of privileged operations, they will only ever use a very limited subset of such functionality during their legitimate execution. For example, the Unix `finger` daemon program executes with root permissions, offering it unlimited privileges. However, during a normal and legitimate run of the daemon program it will only ever use a very small set of privileges: binding port 79, writing to `/etc/log`

and executing `/usr/ucb/finger`. An intruder, however, may be able to manipulate bugs or trapdoors within the program to cause it to use other privileges to compromise the security of the machine. The goal of the UC Davis research is to detect such exploitation of program vulnerabilities.

The essence of the approach lies in specifying, for each security critical program, a *privilege profile* which defines the full set of privileges that the program may make use of during its legitimate execution. A simple language is offered for such definitions, based on a predicate calculus model coupled with a regular expression syntax for specifying patterns of filenames. Profiles expressed in this way are usually concise, defining a very limited domain of privileged activity for the program. By a process of automated translation, these profiles can be used to derive a set of rules which may be applied to an audit trail to detect anomalous use of privilege, that is any behaviour which falls outside of the predefined domain. The rule sets generated from the privilege profiles are typically small and involve little computation, thus permitting real-time analysis of audit information.

Privilege profiles have been constructed for many security critical Unix system programs (e.g., `sendmail`, `fingerd`, `rdist`), permitting the detection of many known attacks which exploit vulnerabilities in these programs. Furthermore, since the privilege profile for a program defines a minimum set of privileges, the generated rule sets are also effective at detecting vulnerabilities which are as yet unknown. In this regard, the specification-based approach improves upon the typical situation found in classical intrusion detection systems. A second advantage of the UC Davis approach is that, assuming correctness in the privilege profiles, the system will never generate false-positives. One drawback of the system, however, is that it is poor at defending programs which legitimately make use of a broad range of privileges. This includes authentication protocols such as `login` and `rlogin` under Unix.

In general, this approach seems to have some merit despite the fact that it may not cope well with certain broad-ranging programs. The principal challenge faced by this type of system is designing a language which can express a program's full range of valid privileges in a concise form which is easy to verify as correct.

Thumbprinting

A second project at UC Davis considers the problems of accurately identifying the perpetrator of an intrusion or intrusion attempt. Such identification is clearly critical to effective response, yet the facilities implemented by current network protocols provide no direct support for such connection tracing. The principal difficulty which must be dealt with in identifying an intruder is that the virtual circuit which links the intruder's machine to the target machine may follow a long and potentially convoluted path (or chain) through several intermediate machines. In tracing such a circuit back to its source, each intermediate machine must determine which of its many incoming connections passed information to the particular outgoing link which perpetrated the intrusion. This matching process often proves to be quite difficult, hampering effective tracing of such circuits.

An approach which offers significant assistance in making such intermediate matchings is that of connection Thumbprinting [22, 23]. Under such a scheme, each machine in a network periodically computes values which summarize the traffic which has travelled

along each of its connections during the preceding interval of time. This thumbprint value, really just a special type of checksum, has the characteristic that it distinguishes a given connection from unrelated connections, but is identical across two connections which form part of the same chain. This property permits easy matching of related incoming and outgoing connections, thereby allowing for improved tracing facilities which have very low probabilities of error.

The particular scheme adopted by the UC Davis researchers accumulates a 24 byte per minute thumbprint on each network connection. Their system makes some effort to cope with complicating factors such as clock skewing, propagation delays, loss of characters (e.g., due to buffer overflows) and packetization variation. The approach has advantages over other, similar, proposals in that it can trace connection chains which leave the domain over which thumbprinting is implemented, and return again from a different point. Situations exist, however, when tracing via thumbprints is not feasible, such as chains in which each link is encrypted differently.

Overall the idea of connection thumbprinting seems worthwhile, offering as it does an increased opportunity for effective response. It must be kept in mind, however, that such a system is still unable to follow a trail beyond the bounds of the network participating in the thumbprinting scheme.

GrIDS

A third project at UC Davis considers the construction of a network-based system which is capable of detecting large-scale intrusions such as sweep attacks, coordinated distributed attacks and worm attacks. These types of threats are characterized by small actions applied to a large number of machines. A classical intrusion detection system running on any single machine would likely not report the occurrence — it is only when the global picture is considered that the scope and nature of the attack becomes clear.

The Graph-based Intrusion Detection System (GrIDS) [21] is an attempt to effectively protect very large networks (e.g., WANs) from these types of threat. The key functionality of the system is the dynamic construction of *network activity graphs* which capture aspects of the current activity of computers on the network plus the traffic which exists between them. Essentially, such graphs describe a causal structure for network activity, aggregated from individual data items collected by point sources on the network (e.g., per-host intrusion detection systems or network sniffers). From the rule-based analysis of such graphs, known types of large-scale attack can be detected. A policy language is supplied for the construction of such rule-sets.

The GrIDS system has been carefully designed to scale up to very large computer networks, an advantage it holds over previous network intrusion systems. The system is inherently hierarchical — host machines are logically grouped into *departments*, each of which may be collected into higher-level departments. Each department is charged with the job of constructing graphs which describe activity within its constituents. Activity which crosses departmental boundaries is passed upwards for further analysis. To keep graphs to a manageable size, such upward propagation is accompanied by an aggregation step in which the activity of an entire department is collapsed into a single node. Using this approach, the designers believe that the system will be capable of scaling to cover networks which span several thousand computers.

The basic premise of a building a scalable system to detect large (or very large) scale coordinated attacks is sound. The principal difficulty faced by such a system is how to effectively abstract information as it is passed up a hierarchy. The GrIDS system mechanism for coping with detail-loss due to such abstraction seems ad hoc and it is difficult to see how it would work effectively, especially when scaled to very large cases. As such, the GrIDS system must be viewed as a first step — it is worth watching to see what future developments occur.

Other UC Davis Research

In addition to these established research projects, newer work at UC Davis is also investigating aspects of intrusion detection relating to network infrastructures [2] and in particular the protection of network resources such as DNS and routers. As yet, no substantial results have been reported from this research.

4.1.2 Stanford Research Institute (SRI)

Similar in focus to the GrIDS research, the EMERALD (Event Monitoring Enabling Responses to Anomalous Live Disturbances) [17] system presently under construction at SRI is an effort to provide intrusion detection and response across very large networks. Like GrIDS, the main thrust of the system is a distributed hierarchical architecture of components. EMERALD supplies a three-level hierarchy: individual hosts run *service analysis* modules which aim to detect misuse of the local resources; hosts are grouped together into *domains* which are monitored for anomalous patterns of use or inter-host traffic. Finally, all domains in the system are subject to *enterprise-wide analysis* which seeks to catch highly distributed forms of attack such as the Internet Worm [20].

A novel feature of the EMERALD architecture is that each of the monitors deployed within the system, irrespective of the level at which they operate, are instances of a single object class. Each such monitor can be said to be composed of three computational elements: a Profiling Engine (which attempts to detect anomalies in a user's pattern of machine use), a Signature Engine (which attempts to recognize known forms of attack) and a Resolver (an expert system which accepts notifications from the other units and from other monitors, and decides upon a response). Monitors within the system communicate with one another via a *subscription* service secured by public key cryptography. This mechanism allows peers at the same level of the hierarchy to exchange information, and also permits monitors to transmit summarized forms of their findings upwards to their parent. The architecture also provides scope for extensibility, providing an API for third-party detection tools to feed information into the EMERALD hierarchy.

The EMERALD system can be viewed as a logical extension of the IDES and NIDES intrusion detection systems previously developed at SRI. The methodology for detection and response (namely a mixture of profile-based analysis and rule-based matching of known attacks) is unchanged, however in place of a monolithic centralized resource for such analysis EMERALD offers a distributed architecture. This has many advantages, most notably the ability to scale the system to large networks by virtue of the distribution of computational load.

EMERALD is purely design at present, with implementation effort only recently commenced. As such the present accounts convey little impression as to how certain critical issues will be overcome. Given that the architecture is primarily a framework under which a hierarchical intrusion detection system can be built, even after the completion of an EMERALD prototype many of these issues will not be confronted until the environment is used to construct a realistic system. It is unlikely that any insightful results will be forthcoming in the short term.

4.1.3 North Carolina State University and the MCNC

The JiNao [8] system, also currently under development, is a tool for detecting intrusions targetting the infrastructure of a network (the routers and switches), and for responding to such threats. The system makes use of profile-based methods found in classical intrusion detection systems, applying them to the infrastructure by substituting routers/switches in place of host machines in the analyses. Thus, the system works by maintaining profiles of past network routing and management requests received from each neighbour, raising an alert when current use deviates from this historical summary. Transitions to improper protocol states — which may indicate suspicious activity — are similarly reported.

Also present within an operating JiNao system are *remote management* entities which oversee the operation of local detection systems running on a group of routers/switches in the network. These management units implement a set of network applications (with protocol based on SNMP) which can probe the status of the constituents and direct them by issuing commands. This hierarchical grouping of units can be extended to higher levels, thus giving the system the potential for scalability. Although beyond the scope of the current research effort, such structure also permits possibilities for detecting distributed attacks such as sweeps or worms.

This is another system on which implementation has only recently begun. Overall it appears to be an application of fairly established technology to a slightly different domain. As such it offers only limited insight into future techniques for intrusion detection, although it highlights the need for such systems at a network infrastructure level.

4.2 Other North American Research

4.2.1 Purdue University

A recent area of research at Purdue's Computer Operations, Audit and Security Technology (COAST) laboratory has involved construction of an intrusion detection system based upon Autonomous Agents [3]. The work is motivated by a desire to overcome limitations present in classical intrusion detections by virtue of their status as monolithic units. These factors, which include lack of scalability and a lack of fault tolerance, can be effectively overcome by considering a system in which intrusions are detected by co-operation between several distributed fine-grain lightweight computational units (i.e., Autonomous Agents). Each agent within the system is highly specialized, being capable of indepen-

dently observing one aspect of the overall environment. Such agents can be added to the system dynamically as required.

One major advantage of such an approach is that the intrusion detection capabilities of the system exhibit a gradual degradation upon failure. If an agent is compromised during an attack, all others may still function effectively (compared to the classical case where a single failure would shut down the intrusion detection system). Furthermore the distributed nature of the computation makes it suitable for scaling to large systems. In the Purdue system, techniques of Genetic Programming are employed to construct agents which are tailored to a particular environment. This can have the effect of making subversion of systems more difficult: knowledge of one agent does not necessarily aid in the subversion of another (since that agent will have been trained slightly differently because it has operated within a slightly different environment).

The arguments for a model incorporating multiple autonomous agents are, on the whole, quite convincing. Also, the idea of "growing" such agents by Genetic Programming seems interesting, although further research would likely be required to determine exactly how to make such a process reliable.

4.2.2 University of New Mexico

The concept of a distributed environment with characteristics of graceful degradation is also embodied by research at the University of New Mexico which considers an immunological approach [6] to intrusion detection. The basis for the research is the analogy of constructing an immune system for the computational environment. Such a system should have properties similar to the human immune system, namely: a protocol of distributed detection in which each detector is unique, a probabilistic model of detection, and the ability to recognize arbitrary foreign objects irrespective of whether they have been previously encountered.

A key to the design of such a system is the definition of self: an immunological detection system can only operate effectively if it can distinguish what elements are part of the host and which are foreign. To this end, a concept of self for a computational process must be developed. The UNM project considers such a definition for Unix processes in terms of sequences of system calls. Security critical programs are observed in a computational environment to determine common system call sequences which occur within their execution. Only short patterns of adjacent calls are considered (e.g., a neighbourhood of six or eleven calls). This information is used to construct a database of signatures, which can be dynamically checked to analyze subsequent usage patterns. Departures from these signatures which go beyond some threshold value are deemed to denote anomalous execution and are reported.

In general, this work is very preliminary, although the concepts which underlie it seem interesting and well founded. The concept of an immune system in which the various detection and response agents can migrate between machines on a network (to concentrate on hot spots) offers an intriguing and tantalizing model of operation.

4.2.3 Texas A&M University

The Cooperating Security Managers (CSM) [24] system prototyped at Texas A&M is another attempt at furthering network intrusion detection in a fashion that is scalable to large networks. Every machine within a network independently runs its own CSM process which is responsible for detecting local intrusions as well as cooperating with other nodes' CSMs to detect large scale attacks. Two schemes are utilized for local detection: a scheme of correlating user activity against a predefined *intruder profile*, plus the recognition of known specific actions known to be suspicious (e.g., an attempt to modify the password file).

Large scale attacks are detected by a mechanism in which a given CSM reports any observed significant activity associated with a network connection to the CSM of the host which originated the connection. This information can be further back-propagated to a point where the extent of the threat becomes known. To assist in response, the CSM system also implements a protocol of user-tracking. Whenever a request for remote access is made to a machine running the system, the requesting machine's CSM coordinates with the target machine CSM. During this negotiation, the requesting CSM transmits a list of all the hosts this user has travelled through to reach this point in the connection chain. The target machine can use this information to respond to any threats which are subsequently discovered originating from this connection.

This research is of some interest, primarily because of its facility to maintain a trail for each user, in preparation for possible attack response. As with all such systems, this function is only useful in following a user's trail between machines within the network of CSM machines. As soon as the trail leaves this network, the information maintained by the protocol is of no assistance.

4.2.4 George Mason University

Research at GMU has considered intrusion detection as an application of a general system of Incremental Inductive Learning [13]. The construction of such a system is viewed as a concept learning problem, characterized by two phases: *start-up* (in which the system is trained by known examples) and *update* (where the system functions autonomously in a live environment and adaptively learns). In the proposed intrusion detection system, both phases centre around the construction and maintenance of use patterns (similar to the profiles collected by classical systems). The learning process involves using feedback from the environment to refine a maintained base of representative examples (both positive and negative). When the system is informed that it has made a mistake, it marks the incorrect decision as a negative example, adds it to the knowledge base and attempts to use this new fact to inductively learn a new minimal set of representative examples.

This work seems largely undeveloped from a practical perspective, focussing primarily on the abstract AI technique rather than its application to computer security. It is of interest only as an account of how machine learning techniques may, in theory, be employed in intrusion detection systems.

4.3 European Research

4.3.1 Supélec Rennes Campus, France

GASSATA (Genetic Algorithm tool for Simplified Security Audit Trail Analysis) [14] is an intrusion detection system prototype which applies genetic algorithms to the task of recognizing known attack patterns in audit trails. It is envisaged that such a system could augment the expert-system-based audit analyses found in classical systems. The approach adopted involves proposing a problem in which a vector summarizing observed events must be mapped back to a vector summarizing how many instances of each known attack method is indicated by this data. At the commencement of the analysis, a population of random solutions is generated. Through successive iterations (generations) each member solution is analyzed for fitness (how well it explains the observations). Those that display poor fitness are killed off, others that display good fitness are chosen to inter-breed to form combined solutions to be carried through to the next generation. By this process, the French project found that the system converged to a solution which summarized the attack suite with high accuracy in 200 generations.

While this result is interesting, the system's ability to recognizing only known forms of attack renders it of limited practical use. Even slight variations of previously seen forms are unlikely to be identified by such an approach.

4.3.2 University of Namur, Belgium

Recent work at the University of Namur in Belgium [15] has sought to integrate a traditional expert-system based intrusion detection system (ASAX) with a configuration analysis tool. This latter system is charged with the task of dynamically scanning the system in order to build up a picture of which sections are (or have become) vulnerable to attack. This knowledge-base can be used to drive the rule-based intrusion detection system, for example informing the expert system which threats it should be looking for in the audit log. While this provides an interesting optimization to classical techniques, it does not represent a major advance in the technology.

4.3.3 University of Technology, Cottbus, Germany

A system for network intrusion detection (AID [19]) has been developed around a simple model in which each host sends auditing information, expressed in an architecture independent language, to a centralized expert system via a secure RPC channel. This architecture is clearly non-scalable and represents only a slight improvement over classical network-wide intrusion detection systems.

4.4 Other Research

To date, no accounts of current U.K. research into the field of intrusion detection have been located.

In the Asia-Pacific region, there are a number of groups who seem to have an interest in researching this domain. University projects at Melbourne University and the University of Woolongong maintain web pages [16, 1] outlining proposed work on such systems. Furthermore, the latter has described a design for an intrusion detection system based around principles of Evidential Reasoning [5]. This is a preliminary mathematical analysis with little practical information.

5 Trends and Conclusions

The research projects we have examined constitute attempts to improve the state of the art in intrusion detection in three principal ways:

1. by providing scalability to large networks,
2. by offering the ability to trace perpetrators along lengthy communication chains, and
3. by investigating alternative AI techniques for the detection of intrusions.

The first of these areas seems an active topic of research, with an implicit consensus that large-scale network-based intrusion detection systems will become critical in the future. It is clear from the DARPA sponsorship of projects such as EMERALD and GrIDS that this is a position strongly held within defence INFOSEC circles. The motivation for such an effort would seem to be a perceived need to protect large networks from co-ordinated attacks of the type envisaged for Information Warfare.

The research into intruder tracing is largely of a preliminary nature, although the techniques being suggested obviously offer some useful facilities for automated response. Response technologies are critical to effective intrusion detection, particularly in situations (such as prevails in the defence arena) where it is essential that security breaches are closed quickly. It seems clear that future systems targetting such sensitive domains must necessarily employ some form of pro-active response centred upon a paradigm of intruder tracing. While Thumbprinting and the simple connection-time scheme of CSM offer some simple tools for such analysis, future research should be directed at investigating schemes which are more generally applicable (e.g., providing some assistance in tracing beyond the local network).

The final area of current study for intrusion detection systems — alternative AI identification techniques — seems relatively broad, with many different AI technologies having been considered for the task. The most attractive of the approaches considered is that which considers the construction of an artificial immune system for the computational environment from a set of Autonomous Agents. This system offers a naturally distributed architecture which avoids the single point of failure/compromise common to most proposals. The genetic approach to identification/training seems to offer several interesting possible advantages which may make it a useful path to pursue further.

Drawing together the body of research surveyed in this report, we can identify a number of technologies that could usefully be incorporated into the next generation of intrusion

detection systems. Firstly, to permit the possibility of scalability to large networks, a system should adhere to a decentralized architecture. This might take the form of a hierarchical (tree-based) scheme, or may simply involve a distributed model of decision-making. Secondly, a system which aims to protect sensitive or classified information should definitely employ some form of automated response capable of closing a security breach quickly. Making such a response pro-active will almost certainly involve a mechanism for tracing an intruder along a potentially convoluted communications chain to the source. Finally, the intrusion detection system should not be susceptible to a single point of failure or compromise: a graceful degradation is preferred.

References

1. Centre for Computer Security Research, University of Woolongong, Web Page: URL http://www.cs.uow.edu.au/ccsr/.

2. Steven Cheung, Karl N. Levitt, and Calvin Ko. Intrusion detection for network infrastructures. Short Presentation at the 1995 IEEE Symposium on Security and Privacy, Oakland, California. Abstract of paper available on-line at http://seclab.cs.ucdavis.edu/arpa/arpa.html, 1995.

3. Mark Crosbie and Gene Spafford. Defending a computer system using autonomous agents. In *Proceedings of the 1995 National Information Systems Security Conference (NISSC95)*, 1995. Available on-line at http://www.cs.purdue.edu/homes/mcrosbie/research/NISSC95/.

4. Dorothy E. Denning. An intrusion detection model. In *Proceedings of the IEEE Symposium on Security and Privacy*, pages 118–131, 1986.

5. Mansour Esmaili, Reihaneh Safavi-Naini, and Josef Pieprzyk. Evidential reasoning in network intrusion detection systems. In *Proceedings of the First Australasian Conference on Information Security and Privacy (ACISP'96), Woolongong*, 1996.

6. Stephanie Forrest, Thomas A. Longstaff, Steven A. Hofmeyr, and Anil Somayaji. A sense of self for unix processes. In *Proceedings of the 1996 Symposium on Security and Privacy, Oakland, California*, pages 120–128, May 1996. Postscript available on-line from http://www.cs.unm.edu/~steveah/research.html.

7. Kathleen A. Jackson, D. DuBois, and Cathy Stallings. An expert system application for network intrusion detection. In *Proceedings of the 14th Department of Energy Computer Security Group Conference*, number LA-UR-91-558, October 1991.

8. Y. Frank Jou, Shyhtsun Felix Wu, Fenamin Gong, W. Rance Cleaveland, and Chandru Sargor. Architecture design of a scalable intrusion detection system for the emerging network infrastructure. Technical Report CDRL A005, Information Tehcnologies Division, MCNC and Department of Computer Science, North Carolina State University, April 1997. Postscript available on-line from http://www.mcnc.org/HTML/ITD/ANR/JiNao.html.

9. Calvin Ko, George Fink, and Karl Levitt. Automated detection of vulnerabilities in privileged programs by execution monitoring. In *Proceedings of the 1994 Computer Security Application Conference*, 1994. Postscript available on-line from http://seclab.cs.ucdavis.edu/arpa/arpa.html.

10. Lawrence Livermore National Laboratory, Sandia National Laboratory. *National InfoSec Technical Baseline: Intrusion Detection and Response (Draft)*, October 1996. Available on-line as http://doe-is.llnl.gov/nitb/ids.html.

11. Teresa F. Lunt. Detecting intruders in computer systems. In *Proceedings of the 1993 Conference on Auditing and Computer Technology*, 1993. Postscript available on-line from ftp://ftp.csl.sri.com/pub/nides/.

12. Teresa F. Lunt, Ann Tamaru, Fred Gilham, R. Jagannathan, Calveh Jalili, and Peter G. Neumann. A real-time intrusion detection expert system (IDES). Technical Report SRI-CSL-92-05, Computer Science Laboratory, SRI International, Feb 1992.

13. Marcus A. Maloof and Ryszard S. Michalski. A method for partial-memory incremental learning and its application to computer intrusion detection. In *Proceedings of the 7th IEEE Conference on Tools with Artifical Intelligence*, pages 392–397, November 1995. Postscript available on-line from http://www.isle.org/~maloof/.

14. Ludovic Mé. GASSATA: a genetic algorithm as an altenative tool for security audit trails analysis. Available on-line at http://www.supelec-rennes.fr/rennes/si/equipe/lme/these/oakland95/oakland95.html, 1995.

15. Abdelaziz Mounji and Baudouin Le Charlier. Detecting breaches in computer security: A pragmatic system with a logic programming flavour. In *Proceedings of the Eighth Benelux Workshop on Logic Programming*, September 1996. Postscript available on-line from http://www.info.fundp.ac.be/~cri/DOCS/asax.html.

16. Obsidian Project Web Page: URL http://obsidian.cs.mu.oz.au/A/indexA.html.

17. Phillip A. Porras and Peter G. Neumann. EMERALD: event monitoring enabling responses to anomalous live disturbances. Technical report, Computer Science Laboratory, SRI International, February 1997. Submitted to the 20th National Information Systems Security Conference; Postscript available on-line from http://www.csl.sri.com/emerald/index.html.

18. Steven R. Snapp, J. Brentano, Gihan V. Dias, T.L. Goan, L. Todd Heberlein, Che-Lin Ho, Karl N. Levitt, and Biswanath Mukherjee. A system for distributed intrusion detection. In *COMPCON Spring '91 Digest of Papers. San Francisco, California*, pages 170–176, 1991. Abstract available on-line as http://seclab.cs.ucdavis.edu/papers/sbd91.abs.

19. M. Sobirey, B. Richter, and H. Koenig. The intrusion detection system aid - architecture and experiences in automated audit analysis. In *Communications and Multimedia Security II (Proc. of the IFIP TC6 / TC11 International Conference on Communications and Multimedia Security CMS'96, Essen, Germany)*, 1996.

20. Eugene Spafford. The Internet Worm: Crisis and aftermath. *Communications of the ACM*, 32(6):678–687, June 1989.

21. S. Staniford-Chen, S. Cheung, R. Crawford, M Dilger, J. Frank, J. Hoagland, K. Levitt, C. Wee, R. Yip, and D. Zerkle. GrIDS — a graph based intrusion detection system for large networks. In *Proceedings of the 1996 National Information Systems Security Conference (NISSC96), Baltimore, Maryland*, 1996. Postscript available on-line from http://seclab.cs.ucdavis.edu/arpa/arpa.html.

22. Stuart Staniford-Chen. Distributed tracing of intruders. Master's thesis, Department of Computer Science, University of California at Davis, 1995. Postscript available on-line from http://seclab.cs.ucdavis.edu/arpa/arpa.html.

23. Stuart Staniford-Chen and Todd Heberlein. Holding intruders accountable on the internet. In *Proceedings of the 1995 IEEE Symposium on Security and Privacy, Oakland, California*, May 1995.

24. Gregory B. White, Eric E. Fisch, and Udo W. Pooch. Cooperating security managers: A peer-based intrusion detection system. *IEEE Network*, 10(1), 1996.

DISTRIBUTION LIST

Directions for Intrusion Detection and Response: A Survey

Dean Engelhardt

	Number of Copies
DEFENCE ORGANISATION	
Task Sponsor	
Director General, R&D	1
S&T Program	
Chief Defence Scientist FAS Science Policy AS Science Corporate Management	1
Director General Science Policy Development	1
Counsellor, Defence Science, London	Doc Data Sht
Counsellor, Defence Science, Washington	Doc Data Sht
Scientific Adviser to MRDC, Thailand	Doc Data Sht
Director General Scientific Advisers and Trials Scientific Adviser Policy and Command	1
Navy Scientific Adviser	Doc Data Sht
Scientific Adviser, Army	Doc Data Sht
Air Force Scientific Adviser	1
Director Trials	1
Aeronautical and Maritime Research Laboratory	
Director, Aeronautical and Maritime Research Laboratory	1
Electronics and Surveillance Research Laboratory	
Director, Electronics and Surveillance Research Laboratory	1
Chief, Information Technology Division	1
Research Leader, Information Technology Division	1
Head, Advanced Computer Capabilities Group	1
Head, C3I Operational Analysis Group	1
Head, CCIS Interoperability Lab	1
Head, Command Support Systems Group	1
Head, Communications Integration Group	1
Head, Human Systems Integration Group	1
Head, Information Architectures Group	1
Head, Information Management and Fusion Group	1
Head, Information Warfare Studies Group	1

Head, Intelligence Systems Group	1
Head, Software Systems Engineering Group	1
Head, Systems Simulation and Assessment Group	1
Head, Trusted Computer Systems Group	1
Author	1
DSTO Libraries	
Library Fishermens Bend	1
Library Maribyrnong	1
Library Salisbury	2
Australian Archives	1
Library, MOD, Pyrmont	Doc Data Sht
Library, MOD, Stirling	Doc Data Sht
Capability Development Division	
Director General Maritime Development	1
Director General Land Development	1
Director General C3I Development	1
Corporate Information Program	
Director General, Information Policy and Plans	1
Director General, Information Strategic Concepts	1
Navy	
SO(Science), Director of Naval Warfare, Maritime Headquarters Annex, Garden Island	Doc Data Sht
Army	
ABCA Office, G-1-34, Russell Offices, Canberra	4
SO(Science), HQ 1 Division, Milpo, Enoggera, Qld 4057	Doc Data Sht
Air Force	
Director, IW	1
Intelligence Program	
DGSTA, Defence Intelligence Organisation	1
Library, Defence Signals Directorate	Doc Data Sht
ASINFOSEC, Defence Signals Directorate	1
Acquisitions Program	
Corporate Support Program(libraries)	
Officer in Charge, TRS, Defence Regional Library, Canberra	1
Officer in Charge, Document Exchange Centre	1

Additional copies for DEC for exchange agreements	
US Defense Technical Information Center	2
UK Defence Research Information Centre	2
Canada Defence Scientific Information Service, Canada	1
NZ Defence Information Centre, New Zealand	1
National Library of Australia	1
B&M Program	
ASSEC, Defence Security Branch	1
ABSTRACTING AND INFORMATION ORGANISATIONS	
INSPEC: Acquisitions Section Institution of Electrical Engineers	1
Library, Chemical Abstracts Reference Service	1
Engineering Societies Library, US	1
Materials Information, Cambridge Science Abstracts	1
Documents Librarian, The Center for Research Libraries, US	1
INFORMATION EXCHANGE AGREEMENT PARTNERS	
Acquisitions Unit, Science Reference and Information Service, UK	1
Library – Exchange Desk, National Institute of Standards and Technology, US	1
SPARES	
DSTO Salisbury Research Library	10
Total number of copies:	68

DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION
DOCUMENT CONTROL DATA

1. CAVEAT/PRIVACY MARKING

2. TITLE

Directions for Intrusion Detection and Response: A Survey

3. SECURITY CLASSIFICATION

Document (U)
Title (U)
Abstract (U)

4. AUTHOR(S)

Dean Engelhardt

5. CORPORATE AUTHOR

Electronics and Surveillance Research Laboratory
PO Box 1500
Salisbury, South Australia, Australia 5108

6a. DSTO NUMBER	6b. AR NUMBER	6c. TYPE OF REPORT	7. DOCUMENT DATE
DSTO–GD–0155	AR-010-350	General Document	October, 1997

8. FILE NUMBER	9. TASK NUMBER	10. SPONSOR	11. No OF PAGES	12. No OF REFS
N9505/13/138	DEF 96/139	DSD	28	24

13. DOWNGRADING / DELIMITING INSTRUCTIONS

Not Applicable

14. RELEASE AUTHORITY

Chief, Information Technology Division

15. SECONDARY RELEASE STATEMENT OF THIS DOCUMENT

Approved For Public Release

OVERSEAS ENQUIRIES OUTSIDE STATED LIMITATIONS SHOULD BE REFERRED THROUGH DOCUMENT EXCHANGE CENTRE, DIS NETWORK OFFICE, DEPT OF DEFENCE, CAMPBELL PARK OFFICES, CANBERRA, ACT 2600

16. DELIBERATE ANNOUNCEMENT

No Limitations

17. CITATION IN OTHER DOCUMENTS

No Limitations

18. DEFTEST DESCRIPTORS

Intrusion Detection
Computer Information Security
Secure Communication
Electronic Security
Computer Security

19. ABSTRACT

This document presents a review of recent research into the automated detection of attack on computers or networks of computers. It is now widely regarded that despite efforts to secure computer systems against intruders by operating system protection, an increasing number of such attempts are succeeding. Only through the careful monitoring of activity in a computation environment can such penetrations be detected and potentially repelled. We briefly describe some early attempts to provide such monitoring and then proceed to describe several research efforts that are currently underway to overcome the limitations of these classical systems. From an analysis of these new approaches we distill several core principles that are critical to the success of future detect and react systems.