Development of an Integrated Solution for Intrusion Detection: A Model Based on Data Correlation

João Afonso
Pedro Nunes Institute
Coimbra, Portugal
jafonso@ipn.pt

Edmundo Monteiro
University of Coimbra
Coimbra, Portugal
edmundo@dei.uc.pt

Vitor Costa
Fisheries Inspection Department
Lisbon, Portugal
vitorcosta@igp.pt

Abstract

This work describes a solution for intrusion detection that presents an improved operational efficacy – both in terms of performance as well as volume of processed data – reducing at the same time the number of false negative and false positive results. For that purpose we correlate the data collected by the intrusion detection system with other data sources, such as events that are reported by interfacing equipment (edge devices) as well as other agents considered crucial for this purpose such as vulnerability detection solutions. As part of the proposed solution the data is collected in a Relational Data base System, to facilitate data correlation, as well as making it available through an easy to use web interface. Additionally, the system interacts with the network managers, in response to pre-defined triggers using a unified messaging platform that uses tools capable of processing E-Mails, Text Messages and also an instant messaging tool based of the XMPP protocol.

1. Introduction

Intrusion Detection Systems (IDS) are becoming increasingly today an essential tool for monitoring safety in computer systems and networks. Their use nowadays is so widespread that more and more appliances have already some sort of IDS mechanism [1] embedded on their systems.

The way these tools operate (either implemented by hardware or software) may vary. Some systems use a variety of remote sensors (network IDS) that execute the task of controlling the packets that travel on a certain point of the network that is considered critical and transmitting relevant information to a central server. Others just have IDS [2] and an agent operating in the same equipment (host IDS). The analysis of the intercepted packets is then confronted with a set of known signatures that, when confirmed, can be considered an intrusion attempt. Nevertheless, an intrusion attempt does not necessarily means an intrusion as we don't know at this stage if the intrusion was succeeded. In order to maintain the efficacy IDS need to periodically update the signature database.. In addition to this set of known signatures IDSs are able to detect a set o situations that may indicate an attack in progress (e.g. Port scanning, etc).

One of the main limitations of IDS solutions is that their are unable to intrusion attacks for which they are not prepared, allowing for an attack to be initiated and even concluded before being detected. This kind of event is classified as a false negative and we estimate that 12% of the IDS records are if fact false negative events. On the other hand, false positive events occur when IDS reports possible intrusions that do not correspond to the reality of the infrastructure to being controlled. This, according to our study, amounts to about 17% of the recorded events. The argument regarding if a certain intrusion is or not correctly classified lies in the fact that an attack on a specific network that is not vulnerable to that type of attack should not be reported (e.g. IDS reporting an intrusion attempt on a Web server using techniques specific for the Microsoft IIS platform when the server is running open source Apache, immune to those attacks).

Several studies are available that address some of the issues discussed here. From those, we highlight solution [3] in which a significant contribution was made towards the reduction of false positive and negative results integrating IDS with a solution for detecting vulnerabilities by correlating data between both.

This paper proposes a solution for intrusion detection that presents an improved operational efficacy, both in terms of performance as well as volume of processed data. The solution proposed aims not only to reduce false negative but also false positive events.

The remaining of the paper is structured as follow. Section 2 provides de system requirements. Section 3 describes the proposed solution. Section 4 presents a case study used for validation of the proposal. In Section 5 the results gathered in the case study scenario are analyzed. Finally, Section 6 presents some conclusions form the present work and some direction for further work.

2. System Requirements

The essential requirements for the intrusion detection system proposed include:

- The integration of IDS with a relational database management system (RDBMS), to allow data correlation with other potential sources, within a desired time frame;
- The possibility of using a Web platform, to monitor the solution real-time. This should also allow simultaneous searching on all data sources;
- Not less important, is the fact that the solution is based on a system for alarm management that is able to notify the person responsible from each occurrence and in able to interact with trouble ticketing solutions.

As it can be seen in Fig. 1, all these requirements should be unified in a single solution capable of allowing access to a wide variety of information through a Web interface.

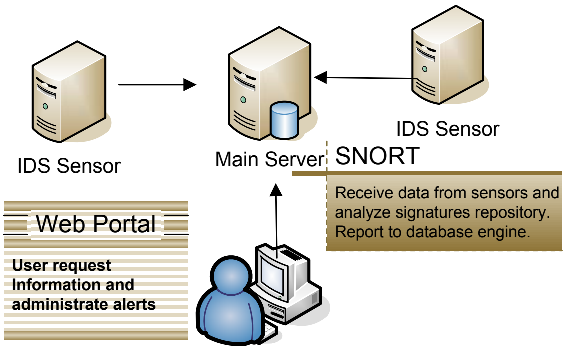

IDS Sensor
Main Server
IDS Sensor
SNORT
Receive data from sensors and analyze signatures repository. Report to database engine.
Web Portal
User request Information and administrate alerts

Figure 1. Data stored in the RDBMS accessible via the WEB interface

2.1. Database engine

The database management system should be able to integrate the IDS and the vulnerability records and should preferably be a free of charge solution to avoid adding unnecessary costs to the solution.

The database engine should support de definition of a list of vulnerabilities according to one of the known repositories: Mitre's Common Vulnerability Enumeration program [4] or the Computer Emergency Response Team's CERT advisories [5].

Considering that an IDS solution has a high event processing load, the RDBMS should be located in a separate server, to optimize the performance of the global solution.

2.2. Web platform

Access to the catalogued data should be made through a Web portal. Also here, preference should be given to free solutions, both for the web server as well as for the development solution.

Through this interface, two levels of operation should be possible: restricted users – with searching and data correlation operations; and administrative users, with privileges to configure all operating components.

The management of alerts should also be operated at this level, not only for characterizing the type of events that should generate alerts but also they way they should be handled (including notifications mechanists and their association with the respective users).

Another functionality supported should be the generation of reports in PDF in order to monitor the functioning of the system.

2.3. Data correlation

Data correlation module should allow different data sources with events collected from its own agents and sensors, to be analyzed in same time period.

Users can view all information, selecting a time windows to analyze, through web portal

The user can check the correlated data from all sources of the RDBMS for any given time period as can be seen in Fig. 2.

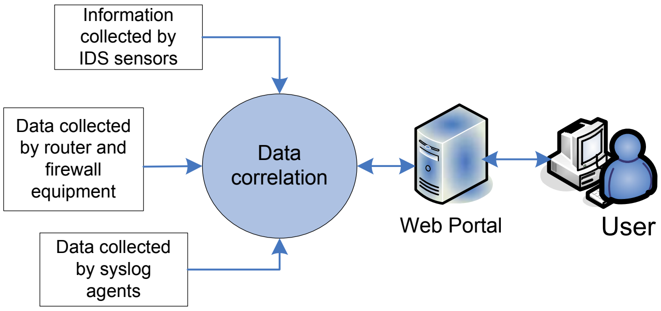

Information collected by IDS sensors
Data collected by router and firewall equipment
Data collected by syslog agents
Data correlation
Web Portal
User

Figure 2. Access from the Web Portal to the correlated data

Globally, the solution should work as represented in Fig. 3. When a potential intruder tries to initiate an attack on a company through the Internet access, the multiple IDS sensors should be able to detect the attack according to a set of known signatures and report to the RDBMS recording all the details associated with the event – date, time, source IP address, destination IP address, source port and destination port. In case of an alarm being activated, the administrator should be notified of the event and should be able, through the web portal, to obtain complementary data about the attack, as well as checking other recent events, to obtain as much relevant information as possible to evaluate the attack in progress.

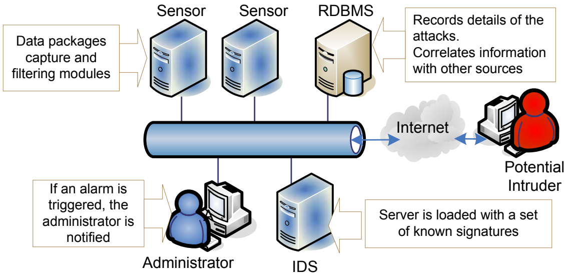

Sensor
Sensor
RDBMS
Data packages capture and filtering modules
Records details of the attacks.
Correlates information with other sources
Internet
Potential Intruder
If an alarm is triggered, the administrator is notified
Server is loaded with a set of known signatures
Administrator
IDS

Figure 3. Global diagram of the solution to implement

3. Proposed solution

The proposed solution was described in detail in a previous work [6] and it includes:

- Operating system: Sun Solaris 10, Intel platform, with MySQLv4 as RDBMS;
- Web server: Apache; with operational development in PHP4;
- IDS: Snort , a popular open source solution [7];
- Vulnerability Assessment Tool: Eeye Retina – commercial solution as a suitable free tool could not be found [8];
- Client data collection service for the RDBMS: Syslog [9];
- Log parsing executed over critical applications [10];
- Instant messaging service, based on the XMPP protocol [11].

The main components of the system can be modeled in a concentric hierarchical form of a circle (Fig. 4) in which the center is the operating system over which is based the data base engine that collects data from the various sources and makes it available through the Web portal.

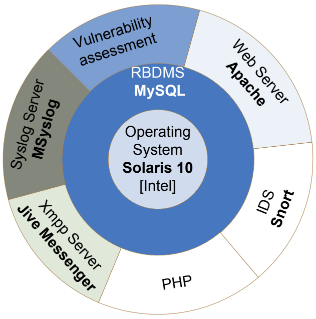


Figure 4. Hierarchical representation of the main components of the solution

For each event processed by IDS, a record is made in the central system. This will be added to the data collected by the Vulnerability Assessment (VA) and the data collected by the agents running on the Windows and Unix systems. Data correlation between of all the sources will then take place using a set of rules that includes time frames, data source/destination, network components involved and also a combination of key field values that signal an already typified attack. In case an event is confirmed, a record is made, together with a score that can be compared with a set value to establish the likelihood that an intrusion attempt has occurred.

As an end result, the number of false positive and false negative results associated with a traditional IDS system is greatly reduced.

Fig. 5 presents a diagram of the proposed solution, centered in the RDBMS that collects data from the various sources and provides the information through the web interface both to restricted users as well as administrators. On the right side of the diagram we can find the alert management system using the three main gateways: SMS, E-Mail and Instant Messaging.

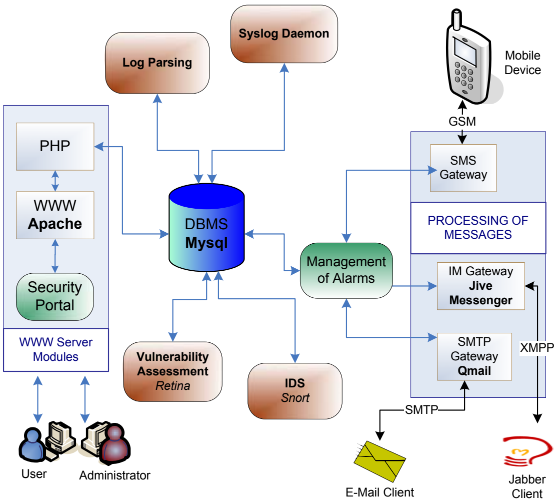
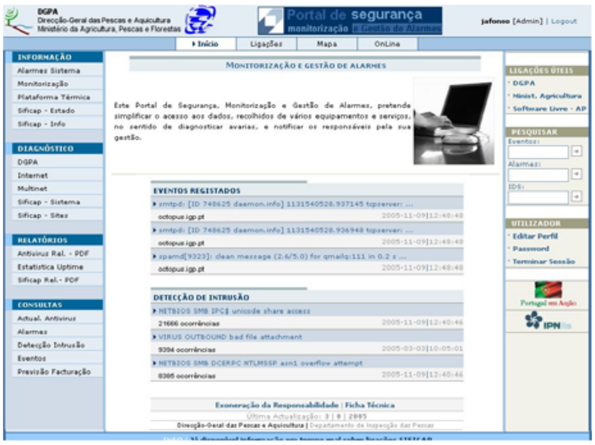


Figure 5. Overview of the proposed solution

4. Case study

The solution presented here was implemented, as a case study, in the General Directory of Fisheries and Aquiculture in Lisbon (Portugal). The system for intrusion detection, as well as the other data sources and the developed web platform was placed on-line in the 1st of January 2005, and the data from the various agents was collected till the 10th of November 2005.

Fig 6 shows a snapshot of the web interface developed.



MONITORIZAÇÃO E GESTÃO DE ALARMES

Este Portal de Segurança, Monitorização e Gestão de Alarmes, pretende simplificar o acesso aos dados, recolhidos de vários equipamentos e serviços, no sentido de diagnosticar avarias, e notificar os responsáveis pela sua gestão.

EVENTOS REGISTADOS

- smtpd: [ID 748625 daemon.info] 1131540528.937145 tcpserver: ... octopus.igp.pt 2005-11-09|12:48:48
- smtpd: [ID 748625 daemon.info] 1131540528.936948 tcpserver: ... octopus.igp.pt 2005-11-09|12:48:48
- spamd[9323]: clean message (2.6/5.0) for qmailq:111 in 0.2 s ... octopus.igp.pt 2005-11-09|12:48:48

DETECÇÃO DE INTRUSÃO

- NETBIOS SMB IPC$ unicode share access 21666 ocorrências 2005-11-09|12:40:46
- VIRUS OUTBOUND bad file attachment 9394 ocorrências 2005-03-03|10:05:01
- NETBIOS SMB DCERPC NTLMSSP asn1 overflow attempt 8385 ocorrências 2005-11-09|12:40:46

Exoneração da Responsabilidade | Ficha Técnica
Última Actualização: 3 | 8 | 2005
Direcção-Geral das Pescas e Aquicultura | Departamento de Inspecção das Pescas

Figure 6. Web Interface

Besides detecting false positives, the solution also includes mechanisms to collect events from crucial applications that would otherwise be impossible to track by other methods (thus minimizing false negative events).

The Web portal allows the users to operate different tasks - according to their individual profile, restricted user or administrator - from running queries over all existing databases as well as managing users and assign persons responsible for each functional area to be notified according to each type of alerts. It is also possible to define for each user the preferred means for communication (IM, SMS and e-Mail) as well as their availability to receive notifications (working hours, during the night or 24x7).

The user is provided with an initial page indicating the global status of the system, indicating the IDS reported attacks with the largest number of occurrences, as well as the most recent events received by Syslog and other information sources used for data correlation.

5. Results

A summary of the results obtained in the period between 1-1-2005 and 10-11-2005, is presented in Table 1 in the form of non-correlated information as obtained by IDS. This table presents, for each type of attack detected, the number of events in the selected period as well as the data of the last event for a particular attack. The information contained in the RDBMS is, of course, much more detailed indicating not only the address and source/destination port as well as many more details on the detected attack. With these elements, it is possible, through the web interface, to search the log of past events with origin on a certain IP address, and evaluate the extension of the attacks as well as their frequency and eventual pattern/trends important in the evaluation of a particular type of attack.

Table 1 - List of attacks recorded between
1st January 2005 and 10th November 2005

ATTACK	N.	LAST
Virus Outbound bad file attachment	9394	05-03-03
NETBIOS SMB DCERPC NTLMSSP asn1 overflow attempt	8249	05-11-08
ICMP L3retriever Ping	6441	05-11-08
ICMP Destination Unreachable Communication Administratively ...	3826	05-03-03
ICMP PING NMAP	2554	05-03-02
ICMP Destination Unreachable Communication with Destination ...	1694	05-03-03
WEB-MISC Invalid HTTP Version String	819	05-03-02
WEB-MISC robots.txt access	783	05-03-03
SMTP MAIL FROM overflow attempt	745	05-03-01
(http_inspect) BARE BYTE UNICODE ENCODING	359	05-03-02
ICMP PING CyberKit 2.2 Windows	319	05-03-02
ICMP Large ICMP Packet	242	05-03-02
WEB-MISC apache directory disclosure attempt	217	05-03-03
SMTP RCPT TO overflow	179	05-03-01
POP3 TOP overflow attempt	174	05-02-25
(http_inspect) NON-RFC HTTP DELIMITER	164	05-03-02
(http_inspect) APACHE WHITESPACE (TAB)	162	05-02-28
POP3 LIST overflow attempt	135	05-02-28
(http_inspect) DOUBLE DECODING ATTACK	107	05-02-22
(http_inspect) OVERSIZE CHUNK ENCODING	99	05-03-01
SMTP Content-Transfer-Encoding overflow attempt	91	05-03-01
(http_inspect) OVERSIZE REQUEST-URI DIRECTORY	68	05-03-01
WEB-FRONTPAGE /_vti_bin/ access	60	05-03-02
WEB-MISC WebDAV search access	50	05-03-01
ICMP Source Quench	48	05-02-02

Table 1 (cont.) - List of attacks recorded between 1st January 2005 and 10th November 2005

ICMP PING speedera	31	05-01-13
WEB-CGI formmail access	26	05-02-28
WEB-MISC Chunked-Encoding transfer attempt	23	05-03-01
WEB-FRONTPAGE rad fp30reg.dll access	20	05-03-01
WEB-IIS ISAPI .ida attempt	19	05-02-10
WEB-IIS ISAPI .ida access	19	05-02-10 18:24:58
WEB-MISC http directory traversal	18	05-03-03
WEB-IIS view source via translate header	14	05-02-28
ICMP superscan echo	11	05-02-16
WEB-CGI formmail arbitrary command execution attempt	8	05-01-26
DNS zone transfer TCP	6	05-02-20
Attack-Responses 403 Forbidden	5	05-02-25
Web-Frontpage shtml.exe access	4	05-01-13
Web-Frontpage _vti_rpc access	4	05-01-13
WEB-IIS %2E-asp access	4	05-01-20
WEB-MISC weblogic/tomcat .jsp view source attempt	3	05-02-16
POP3 DELE overflow attempt	3	05-01-25
DNS named version attempt	3	05-02-13
SMTP chameleon overflow	3	05-02-01
WEB-CLIENT Microsoft emf metafile access	3	05-01-10
WEB-IIS cmd.exe access	2	05-01-03
WEB-MISC cross site scripting attempt	1	05-02-12
WEB-IIS encoding access	1	05-02-17

From the information included in Table 1 we can identify several attempts to attack the system that must be signaled as false positive because they are not correlated with vulnerabilities that apply to the infrastructure, such as:

- ICMP PING CyberKit 2.2 Windows – indicates possible scanning from hosts infected with Welchia/Nachi worm, with 319 occurrences;
- WEB-FRONTPAGE /_vti_bin/ access – only applies to Microsoft Internet Information Server (IIS), with 60 occurrences.

To verify the situation described data is automatically correlated with existing information in the RDBMS supplied by the VA tool, using as foreign key the CVE references that must be unique.

For the detection of false negatives data is analyzed from the information collected by Syslog – also centralized in the RDBMS – as well as other information sources concerning critical applications. A specific example of a false negative result that will be detected are undue accesses to the Listener port in Oracle that, in case it is not protected by password, will allow a set of basic commands to be executed that might generate events, that would otherwise be undetected by IDS.

As an example we present in Fig. 7 the evolution of the attacks registered in the period between 19th October and 9th November 2005, with a two day sampling interval, where it can be observed that the number of attacks per day remains practicably constant. The initial value of the graph refers the first day of recording because it does not occur at time 00:0h. With this exception, we can verify daily rates of attacks between 600 and 650 per day.

It is also interesting to note that, despite the updates of the SNORT rules, no significant fluctuations on the number of daily attacks was reported.

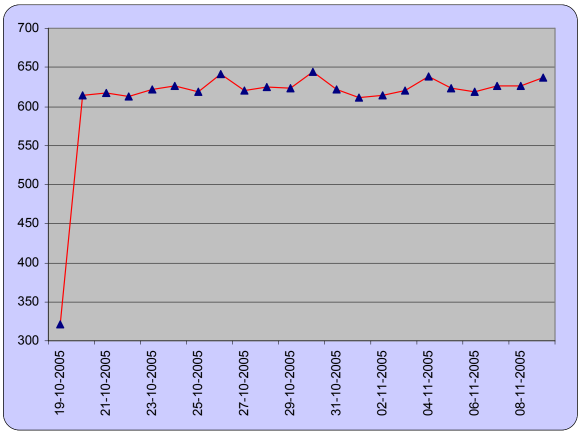


Figure 7. Number of daily events between 19th of October and 9th of November 2005

Analyzing the number of attacks per week day (Fig. 9) we can observe that Wednesdays and Thursdays register the largest number of hits while weekends have a somehow lower number of occurrences.

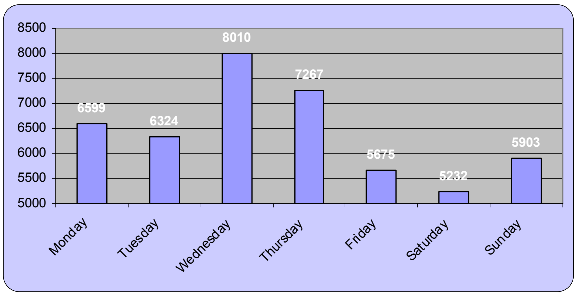


Figure 8. Number of registered attacks by weekday

The back log regarding the source of each one of the attacks is kept in a separate table of the RDBMS registering the date and type of the attacks registered according to their origin. This type of information can be correlated wit other external systems permitting a global record of potential intruders.

6. Conclusions and Future Work

The results obtained with the work described in this paper allow to conclude that it is possible to reduce substantially the number of false positive events as well as to create new intrusion detection mechanisms from potential attacks that might otherwise be undetected using a solution based on data correlation between different sources.

Overall, we consider the solution proposed to be:

- More reliable that traditional IDS solutions, by allowing a better surveillance of systems and network security in which it operates;
- More robust, because it does not overload the RDBMS with futile events that are not in accordance with the records of the VA system;
- More accessible, by allowing an easy access to the data either by the use of the web interface, either by extrapolating from the extensive results obtained by the sources that feed the RDBMS;
- More efficient, by integrating trouble ticketing management tools that allow an easy notification of the person or group responsible for each area, through SMS, E-Mail and/or instant messaging.

As a proposal for future development lies the challenge of improving the capacity to further reduce false positives by incorporating other VA tools, such as Nessus, among others. The improvement in the detection of false negatives might also include data correlation with other agents. It is increasingly common to find proprietary applications in the clients that generate their own operating events that could be catalogued to provide evidence of potential attacks impossible to detect by any other means.

Another potential further development is outlined in Fig. 9, whereby the users of the system could use a mobile terminal or a PDA, not only to receive notifications of attacks (as they do now) but also to interact with the central system, inhibiting or activating a certain service or system in order to contain or counteract an ongoing attack.

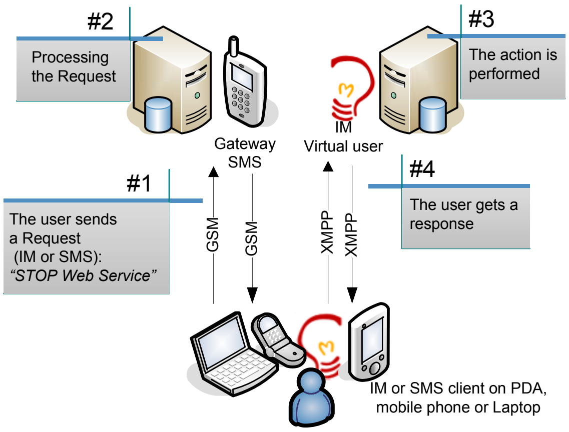


Figure 9. Possibility of users with mobile terminals to interact with the central system

7. References

[1] Nail Desai, IDS Correlation of VA Data and IDS Alerts, Security Focus, June, 2003.

[2] Caswell, B., Beale, J., Foster, J., Faircloth, J., Snort 2.0 Intrusion Detection, Syngress, 2003.

[3] Ron Gula, Correlating IDS Alerts with Vulnerability Information, Publishing House, 2003.

[4] http://www.cve.mitre.org

[5] http://www.cert.org

[6] Afonso, J., Monteiro, E., Ferreira, C., Monitoring and Alarm Management for system and network security: A web-based comprehensive approach, 2nd International Conference on E-Business and Telecommunication Networks, Reading U.K., October, 3-7, 2005.

[7] Gerg, C., Cox, K., Managing Security with Snort and IDS Tools, O'Reilly, 2004.

[8] http://www.eeye.com

[9] Lonvick, C., The BSD syslog Protocol. RFC 3164.

[10] Shearer, A., Firewall log analysis using ACID, SANS Institute, 2003.

[11] P. Saint-Andre, Ed., Extensible Messaging and Presence Protocol (XMPP): Core, RFC 3920, 2004.